

Morphological Segmentation of Blood Images for Automated Malaria Diagnosis

Florian Biermann
`fbie@itu.dk`

November 26, 2013

Contents

1	Introduction	4
1.1	Malaria Diagnosis	4
1.2	Problem Analysis and Report Contributions	4
1.3	Mathematical Morphology	6
2	Related Work	6
2.1	Mathematical Morphology in Automated Malaria Diagnosis	7
2.2	Other Malaria Diagnosis Techniques	7
2.3	General Segmentation of Red Blood Cells	8
3	Morphology in Blood Cell Segmentation	9
3.1	Morphological Primer	9
3.1.1	Connected Set Opening and Closing	10
3.1.2	Area Opening	11
3.1.3	Attribute Opening and Thinning	13
3.2	Estimating Size Distribution of Red Blood Cells	13
3.2.1	Granulometry	13
3.2.2	Area Granulometry	15
3.2.3	Attribute and Shape-only Granulometry	15
4	Ultimate Attribute Opening of Red Blood Cell Images	17
4.1	Ultimate Attribute Opening	19
4.2	Ultimate Attribute Opening on RGB Color Channels	22
5	Implementation	23
5.1	Attribute Opening and Closing	26
5.2	Attribute Granulometry	27
6	Discussion	27
6.1	Analysis of Segmentation Performance	27
6.2	Future Work	30
6.2.1	Improvements	30
6.2.2	Evaluations	31
7	Conclusion	32
A	Segmentation Results	36

List of Figures

1	Examples of images of Giesma stained slides of red blood cells.	5
2	Gray-scale closings of an image of red blood cells.	9
3	Simplified example of gray-scale area opening.	11
4	Increasing gray-scale area closings of an image of red blood cells.	12
5	Normalized area granulometries for images of red blood cells.	14
6	Normalized bounding box aspect ratio granulometries.	16
7	Normalized bounding box fill ratio granulometries.	18
8	Naive foreground model.	19
9	Step-wise computation of foreground model.	20
10	Segmentation for $\epsilon = \frac{\lambda}{2}$ using area closing.	21
11	Separate color channels of the same image.	23
12	Separate color channel segmentation.	24
13	Continued from figure 12	25
14	Segmentation results for separated cells	37
15	Segmentation results for separated cells (continued)	38
16	Segmentation results for separated cells (continued)	39
17	Segmentation results for overlapping cells	40
18	Segmentation results for overlapping cells (continued)	41
19	Segmentation results for overlapping cells (continued)	42
20	Segmentation results for squashed cells	43
21	Segmentation results for squashed cells (continued)	44
22	Segmentation results for squashed cells (continued)	45
23	Segmentation results for blurred cells	46
24	Segmentation results for blurred cells (continued)	47
25	Segmentation results for blurred cells (continued)	48

1 Introduction

This report is concerned with computer vision for diagnosing Malaria disease. Malaria is one of the deadliest diseases world-wide. Diagnosis is usually performed manually, by screening blood samples of patients using light microscopes [38, 15].

Manual diagnosis is a time consuming task that requires special training [38, 15]. Automating the diagnosis of Malaria using computer vision is an alternative to manual diagnosis that can be faster and more accessible.

From a computer vision point of view, the first step in analysis is to separate the image into regions of interest. This process is called segmentation [26]. Regions of interest in this case are single red blood cells. In this report, we will concern ourselves solely with the segmentation step of the computer vision image analysis of malarial blood samples.

1.1 Malaria Diagnosis

Malaria is a disease caused by parasites, which infect human red blood cells (erythrocytes). The parasites are transmitted by the bite of female Anopheles mosquitoes. While treatable if diagnosed early enough, Malaria is still a highly lethal condition [38, 31].

The “gold standard” in Malaria diagnosis today is still the manual counting of infected cells under a light microscope. In order to do so, a blood sample of the patient, usually taken from a finger, is spread out on a thin glass slide and stained with Giesma stain. This is called a thin blood smear sample [38]. Thick blood smears are used for easier detection of Malaria parasites in the blood and not covered in this report. Giesma stain binds to certain proteins in parasites and colors them violet. It also colors red blood cells slightly blue and it also colors white blood cells and smaller artifacts in the blood [15, 38, 29].

The World Health Organization’s (WHO) training manual for Malaria diagnosis recommends to analyze 100 independent fields (i.e. the area that can be seen under the microscope) before giving a negative diagnose and 100 more, as soon as a parasite is found. The 100 extra fields are required to calculate the parasitemia, which is the ratio of infected to uninfected cells. This is a good indicator of the severity of the infection [38, 31].

1.2 Problem Analysis and Report Contributions

From the Malaria diagnosis routine as recommended by the WHO it is clear that a system for diagnosing Malaria must be able to count the single red blood cells correctly. We require therefore a segmentation technique for images of red blood cells with a high sensitivity.

Microscopy images of Malaria slides can exhibit a variety of conditions. Figure 1¹ shows four different conditions of slides. Most notably, red blood cells

¹All images of blood cells are courtesy of F. B. Tek, source: <http://boraytek.blogspot.dk/p/data.html>, accessed on 16.09.2013.

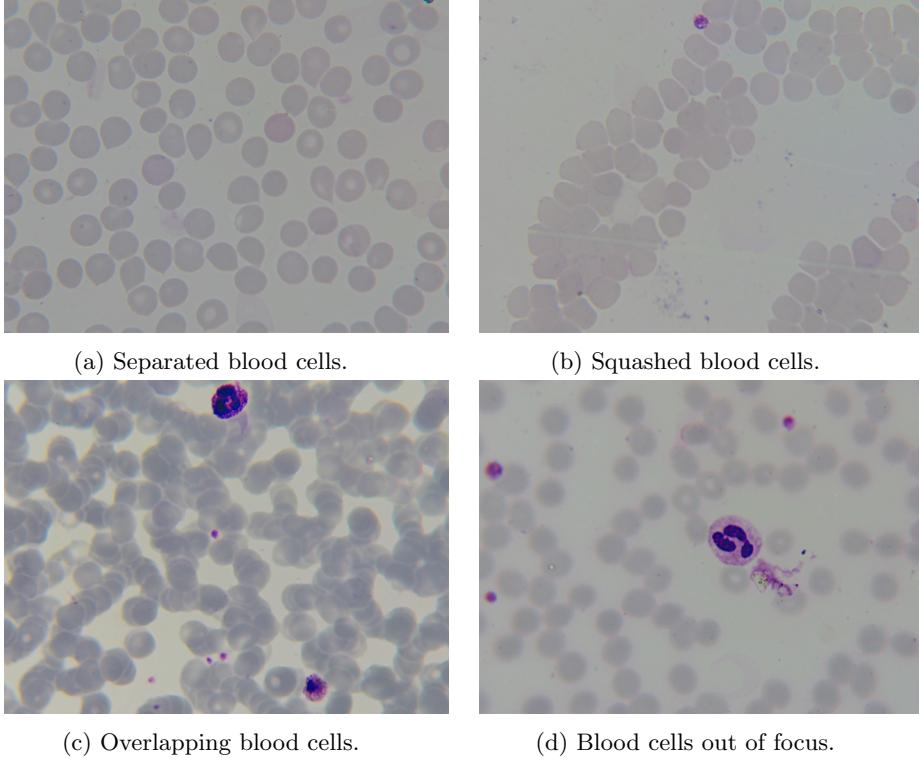


Figure 1: Examples of images of Giesma stained slides of red blood cells.

can be:

- a) Separated
- b) Squashed
- c) Overlapping
- d) Out of focus

Even for trained experts, it can be very challenging to diagnose these kind of blood slides. However, in laboratories, the quality of the slides is usually very high. Bad quality slides are mostly a result of untrained personell preparing these.

The main contribution of this report is to develop and analyze a segmentation approach, based on *ultimate attribute opening* [3, 12, 13] for separate *RGB* (red, green and blue) color channels. Ultimate attribute opening computes the opening of a gray-scale image for which the gray-scale difference is the highest amongst all possible openings. It is based on advanced morphological filtering operations on binary gray-scale images. Advantages of ultimate attribute

opening is that it is essentially parameter free and that it can be computed by applying only two opening operations to the image.

Couple-charged device cameras nowadays provide us with *RGB* color images. For morphological filtering of light-microscopy images, it is crucial to decide how the image should be converted to gray-scale. A common approach is to convert the image into *HSV* color space and to then only select the value channel. This is basically an average of all three *RGB* channels.

However, due to the magnification of images taken through light-microscopes (and also the general use of optical lenses), chromatic aberration occurs. Chromatic aberration is the effect of light of different wave lengths (i.e. colors) focusing at different points [10].

By analyzing the single channels of the *RGB* input image, we can select the color channels which produce the best segmentation results and use these instead of an average gray-scale image, where the channels shadow information of each other.

A side effect of this report is the open-source implementation of morphological attribute opening and closing in C++. These filters are not part of standard imaging libraries available (see section 5).

1.3 Mathematical Morphology

Mathematical morphology is the study of shapes or connected components on binary or gray-scale images [25]. Techniques from mathematical morphology have been used in medical image analysis and especially in automated Malaria diagnosis for a while already [35, 6, 20, 8, 30].

With mathematical morphology it is easy to filter elements from an image that fulfill certain attributes. Morphological Openings and Closings are effective tools for isolating arbitrary, parameterized shapes [5, 33, 32, 3, 12, 13]. Basic morphological operations, erosion, dilation, opening and closing, use structuring elements that alter the structure of an image during filtering. More advanced morphological filters can avoid this problem, and therefore are much more suitable for segmentation purposes.

2 Related Work

While not very popular in research [28, 29], the segmentation of red blood cells, or erythrocytes, for automated Malaria diagnosis is not an unstudied field. In this section, we will take a look at publications that report on previous work on this very topic, both related to automated Malaria diagnosis and seen as a general task.

2.1 Mathematical Morphology in Automated Malaria Diagnosis

DiRuberto et al. [6, 7, 8] propose to use morphological filters to analyze the size distribution of the elements on a gray-scale image of red blood cells. They compute the granulometry, or pattern spectrum [17], using connected set openings for a circular structuring element with increasing radius (see section 3.2.1). Granulometry peaks at the average size of the red blood cells in the image [7]. This estimated average size can be used to improve the segmentation, as the size of the regions we are looking for is known.

However, in [20] Rao and Dempster propose the use of a more general operator, namely area opening [35], to compute the granulometry. Area opening requires no a priori knowledge of the shape of the elements to be filtered. They show that area granulometry outperforms granulometry in finding the average size of the elements in the image, because it deals better with non-uniform shapes, such as cells with holes.

Area granulometry has gained a lot of recognition in research on automated Malaria diagnosis. Tek et al. also used area granulometry to estimate the average size of red blood cells on an image. They use the gained knowledge to compute a segmentation using minimum-area watershed transform (*MAWT*) [28, 30].

MAWT is a watershed algorithm, that uses no mask as seeds. Instead, its only additional parameter is a lower bound on the area of a watershed. Tek et al. report on satisfying segmentation results. However, *MAWT* segmentation is not sensitive enough, so a compound approach of area granulometry, *MAWT*, radon transform and classical watershed is used to perform a segmentation [28], which is complicated and performs slowly.

Moon et al. also use area granulometry as a size-estimator in their study. However, they propose different techniques for actual segmentation of images. Since their method is targeted toward differently stained samples, which are processed in a highly controlled environment only. Also, their goal is not diagnosis but parasite viability detection for Malaria drug development [21].

Instead of using morphological filters for segmentation, they build a foreground model of the image by thresholding. In the subsequent step, they simply divide the area of the foreground by the average cell area and thereby estimate the number of cells present on the image [21]. In subsequent steps they compute the concavity of shapes and to separate touching cells.

2.2 Other Malaria Diagnosis Techniques

Tek et al. contributed beyond the usage of mathematical morphology. Apart from reviewing most literature covering automated Malaria diagnosis [29], they proposed a technique for color-normalization of Giesma stained blood slides [27]. This is helpful for using color-based parasite classification in less controlled environments, where the lightning or the staining intensity can vary greatly.

A study on Malaria parasite detection and classification by Ross et al. reports on the usage of multiple thresholding to segment red blood cells [24]. Ross et al. acknowledge that their segmentation method is not very elaborate and that it can be improved upon. However, their classification produces very specific results [24].

Makkapati and Rao report on an *HSV* color space based thresholding technique for segmenting Leishman-stained samples of red blood cells [14]. Their method computes optimal thresholds for the hue and saturation channels and they report on a very high sensitivity and specificity.

A different, color-based segmentation method by Mandal et al. uses normalized graph-cuts on color images to separate stained parasites from their surroundings [16]. Mandal et al. evaluate their method in a number of color spaces. However, they do not concern their selves with segmenting blood cells.

Other publications do not focus on red blood cell segmentation as much, such as e.g. Anggraini et al., who use Otsu thresholding, [1] or Halim et al. [11]. Instead they investigate techniques for classifying Malaria parasites correctly.

2.3 General Segmentation of Red Blood Cells

Most literature on image segmentation of human red blood cells handles the segmentation of leukocytes, i.e. white blood cells [28]. Still, some research has been conducted on the general task of cell segmentation.

An early publication by Wermser et al. proposes adaptive thresholding of blood images for segmentation [36]. Their study is mainly focused towards finding an approach that is easy to implement and performs quickly. Also, they propose to implement their method in hardware for faster execution.

Ritter and Cooper report on a segmentation method that automatically selects a threshold using connected components that are moved over the image [23]. Their approach is targeted at classifying diseases that affect the shapes of human red blood cells, so they require the borders of the cells to be extracted correctly. A sweep-line algorithm detects the border of each cell after thresholding. Finally, they use Dijkstra's shortest path algorithm on a graph-representation of each border to optimize what is extracted as a cell. Morphology plays a minor role in their method, as they use morphological closings to separate cells that are touching. Their approach deliberately does not deal with overlapping cells [23].

Bergen et al. present a study on the segmentation of white blood cells in blood images [2]. Segmenting red blood cells is to some extend a side-effect of the method they use. Information about red blood cells in the image can improve the segmentation of white blood cells. Using a naive Bayes classifier, they build a probability image of red blood cell locations on the image. A synthetic template, which is a rotated one dimensional Gaussian distribution, is matched to this image. This produces local peaks on the probability image, which represent the centers of red blood cells in the image [2]. Since this is enough information for their purpose, they stop handling red blood cells at this point.

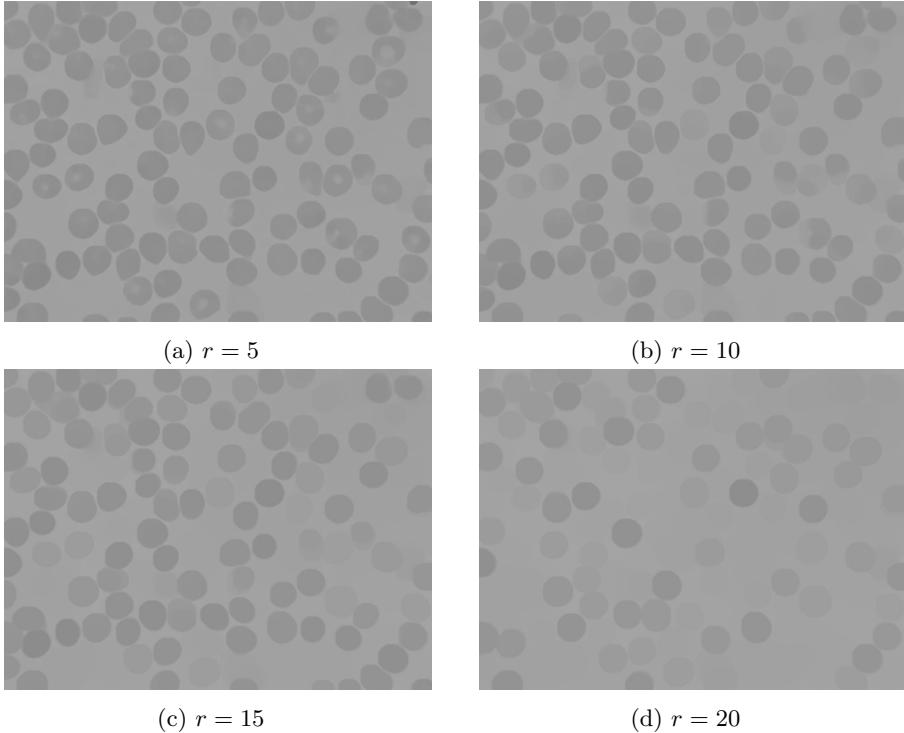


Figure 2: Gray-scale closings of an image of red blood cells using a disk-shaped structuring element with radius r .

3 Morphology in Blood Cell Segmentation

3.1 Morphological Primer

Our approach focuses on using morphological operators to segment images of red blood cells. It is useful to recapitulate mathematical morphology in order to understand the algorithms we use.

Mathematical morphology is the study of shapes in binary or gray-scale images. In our case, these images are two-dimensional, but mathematical morphology can also be applied in higher dimensions [25]. This section is not meant as a comprehensive review of this topic, but to give an overview over the techniques we use.

In this section we will take a look at connected set openings and closings with structuring elements, area openings and the generalization of those, namely attribute openings and thinnings as well as size distributions of shapes over binary and gray-scale images.

3.1.1 Connected Set Opening and Closing

Operations on connected sets require first to define the connectivity between elements. In the case of mathematical morphology, elements are pixels in a gray-scale image, which is the set. Pixels are either 4- or 8-connected. Throughout this report, we assume that pixels are 8-connected, i.e. each pixel is directly connected to all its neighbors.

Basic morphological operators on binary or gray-scale images for a structuring element S are erosion, ϵ_S , and dilation, δ_S . We will only concern ourselves with morphological gray-scale operators in this report.

During dilation, the structuring element S is moved around the outside border of a connected component, or shape, C on the image. S can only move through areas where it does not violate the borders of any shapes. Therefore, it can connect formerly unconnected shapes. The border of $\delta_S(C)$, respectively $\epsilon_S(C)$, is the path of the center of S during dilation. The shape of S has an immediate effect on the result of the operation and changes the structure of the image to reflect S [25]. During erosion, S is moved on the inside border of C . Intuitively, dilation increases the size of a dark element in an image, while erosion decreases it [26].

Erosion and dilation are duals. Mathematically, these are projections $T \rightarrow T$, where T is a lattice, which commute with the inferior (\inf) or superior (\sup) of T [25].

Let $I : M \rightarrow \mathbb{Z}$ be a gray-scale image (M is the mask containing connected sets) and $x \in M$ a point on M [25]:

$$\delta_S(I)(x) = \sup\{\delta(t) \mid t = \begin{cases} z < I(x) & \text{if } \exists x \in T_h(I) \\ -\infty & \text{otherwise} \end{cases}\}, \quad z \in \mathbb{Z} \quad (1)$$

where T_h is a threshold function at $h \in \mathbb{Z}$:

$$T_h(I) = \{x \in M \mid I(x) \geq h\} \quad (2)$$

The dual, ϵ_S , is defined analogously using the \inf .

Morphological gray-scale opening, γ_S , and closing, φ_S , are defined as concatenations of dilation and erosion [25]:

$$\gamma_S = \epsilon_S \circ \delta_S \quad \varphi_S = \delta_S \circ \epsilon_S \quad (3)$$

where \circ is the function concatenation operator. The naming of the operators is very intuitive. A closing “closes” dark holes on a white images, i.e. it removes dark elements. An opening of an image “opens” the dark regions in it, removing bright areas [26].

In figure 2, we see the effects of morphological closing on an image of red blood cells using a disk-shaped structuring element. Figure 2a starts with $r = 5$, r is successively increased by 5 over the following images. Closing successfully filters artifacts and blood cells, but from figure 2c and 2d it becomes clear that it does not preserve the structure of the elements that are not removed.

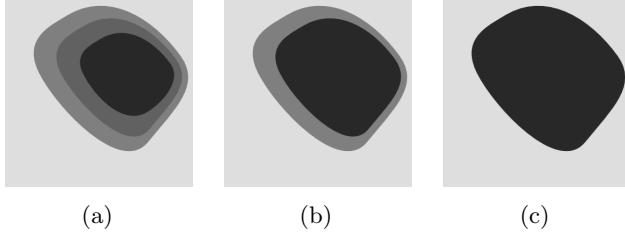


Figure 3: Simplified example of gray-scale area opening.

While still the structuring element S is moved in- or outside of a shape C , it is now the most inner or most outer point of the border of S that determines the shape of $\gamma_S(C)$, $\varphi_S(C)$ respectively [25].

3.1.2 Area Opening

Vincent described the notion of morphological area openings and closings [35]. Area opening and closing are denoted as γ_λ^a and φ_λ^a respectively. Other than morphological openings, morphological area openings do not rely on a structuring element to filter the image. Instead, area opening and closing operate on an area parameter, λ [35].

λ defines the lower area limit of connected components. Every component smaller than λ will be removed. This enables us to filter arbitrarily shaped elements from the image, only based on their sizes. The absolute shape of connected components is only a matter of the image's structure [35]. This means, that we do not require to know anything about the elements we want to filter (except their size of course), while still preserving the shapes in the image while filtering.

Let again $I : M \rightarrow \mathbb{Z}$ be a gray-scale image and let $\|\cdot\|$ denote set cardinality. Gray-scale area opening for a point $x \in M$ is defined as follows [35]:

$$\gamma_\lambda^a(I)(x) = \sup\{h \leq I(x) \mid \|\gamma_x(T_h(I))\| \geq \lambda\} \quad (4)$$

The dual, φ_λ^a , is defined analogously using the *inf*. The most naive algorithm for area opening is to lower a threshold around each regional maximum [34] until its area reaches λ [35]. For area closing, the threshold function is simply inverted.

Figure 3 shows the step-wise computation of area opening on a synthetic image. Figure 3a is the original image, figure 3b is the opening of the original image for a λ greater than the central dark area but less than its directly surrounding area. λ is then increased for each successive image. This illustrates the notion of thresholding for area opening and closing.

In figure 4 we see the effects of morphological area closing on a real image of red blood cells. The blood cells are dark on a light background, so they are removed successively with increasing λ . Figure 4a shows the closing of the

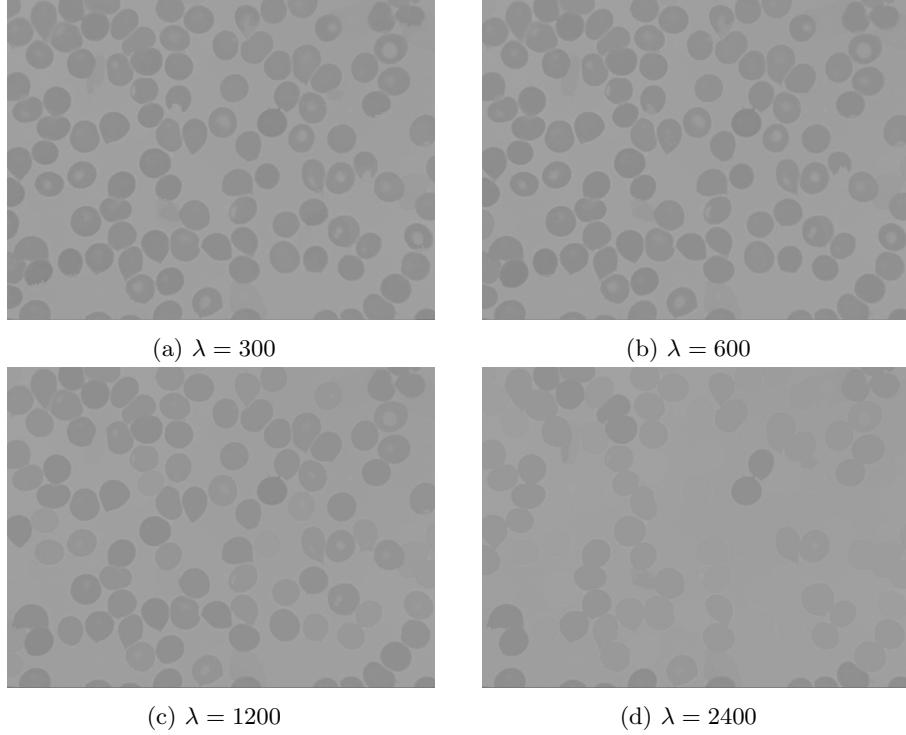


Figure 4: Increasing gray-scale area closings of an image of red blood cells.

image for $\lambda = 300$, figure 4b through 4d show the closing for $\lambda = 600, 1200$ and 2400 , respectively.

Vincent showed that area opening and closing are increasing operators. This means that a connected component enclosing another smaller connected component of lower gray level can never have an area less than the enclosed component. Formally, for each connected set X which has a size greater or equal to λ , its super-sets also must fulfill this condition [35]:

$$X \subseteq Y \Rightarrow \gamma_\lambda^a(X) \subseteq \gamma_\lambda^a(Y) \quad (5)$$

He proposed to use area opening and closing for removing unwanted artifacts from images or to find small elements in an image by first closing or opening the image with some λ and then analyzing the pixel-wise difference of the original and the filtered image [35].

A number of algorithms for area opening and closing have been proposed. However, naive approaches, like the mentioned threshold lowering, can yield very high execution times. Meijster and Wilkinson developed an efficient algorithm based on the union-find data structure [19], which we will use in this report.

3.1.3 Attribute Opening and Thinning

A generalization of area opening and closing is the notion of attribute opening and closing. A further generalization is attribute thinning, which uses a non-increasing attribute [5]. For thinnings, super-sets of any connected set do not necessarily fulfill the same criterion as any subset of them.

The trivial attribute opening of a connected set $C \subseteq M$ for the attribute T is defined as [5]:

$$\Gamma_T(C) = \begin{cases} C & \text{if } C \text{ satisfies } T \\ \emptyset & \text{otherwise.} \end{cases} \quad (6)$$

Substituting the trivial opening of a connected set from equation 6 into equation 4, we get the definition of *attribute opening* [5]:

$$\gamma_T^t(I)(x) = \sup\{h \leq I(x) \mid h \in \Gamma_T(T_h(I))\} \quad (7)$$

Gray-scale closing φ_T^t is, again, defined analogously using \inf [5]. This generalization can be used to perform openings and closings on images using parameterized shape descriptions.

For an attribute T' that is non-increasing, i.e. for which a generalized version of equation 5 does not hold, opening performs a thinning, closing a thickening, respectively [5].

An efficient algorithm for computing attribute openings and closings and granulometries has also been proposed by Meijster and Wilkinson [19, 37].

3.2 Estimating Size Distribution of Red Blood Cells

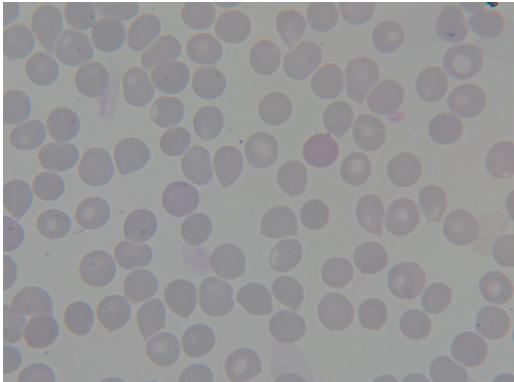
3.2.1 Granulometry

The granulometry or pattern spectrum of a gray-scale image is defined as a differential function of the size of a structuring element for opening or closing of an image to gray value [17]:

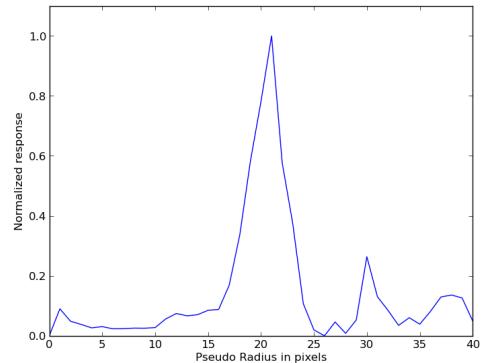
$$g(k) = \sum \gamma_{S_k}(I) - \sum \gamma_{S_{k-1}}(I), k \in \mathbb{N}^+ \quad (8)$$

where I is the image, S_k a structuring element of size k and \sum the sum of gray-scale values of an image.

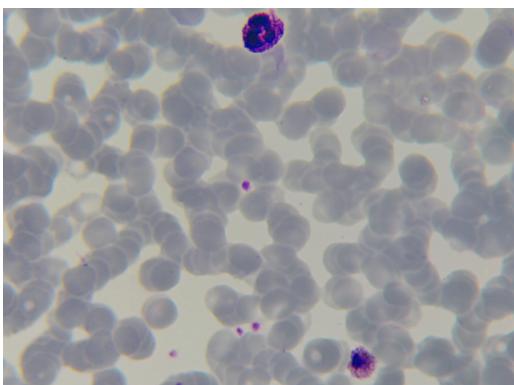
DiRuberto et al. used granulometry with a disk-shaped structuring element to compute a size distribution over the red blood cells in the image [7]. For a disk-shaped structuring element, k is the diameter of the disk in pixels. The central idea is that, since we are looking at an image of red blood cells, the red blood cells should represent the majority on the image and therefore, the granulometry should peak at a k which is this average diameter of a red blood cell in the image.



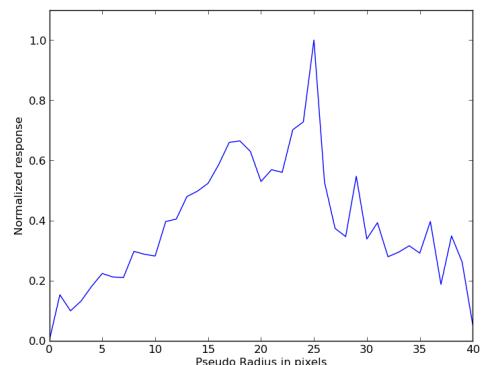
(a)



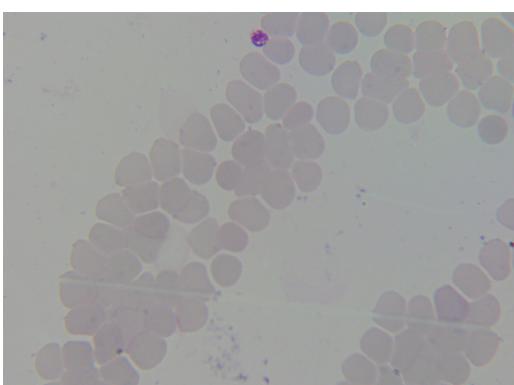
(b)



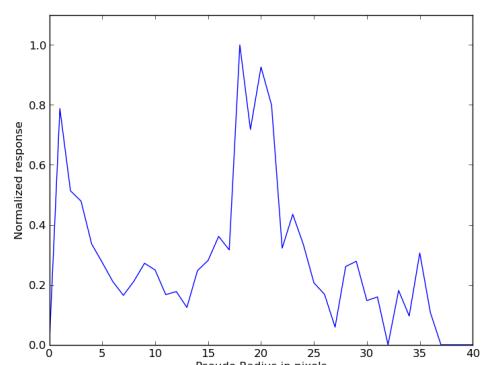
(c)



(d)



(e)



(f)

Figure 5: Normalized area granulometries for images of red blood cells.

3.2.2 Area Granulometry

Red blood cells are, however, not perfectly shaped circles and vary in shape, may have holes and so on. In order to deal with these properties of cells, Mohana and Rao proposed the use of area-granulometry, using the main property of area opening as argument as presented in section 3.1.2: *morphological area-openings require no a priori knowledge of the element we want to filter*. Area granulometry is a more general operator and may be better suited for estimating the average size of red blood cells on an image [20].

We define a function g^a on pseudo-radius to compute the area-granulometry [20]:

$$g^a(k) = \sum \gamma_{\pi k^2}^a(I) - \sum \gamma_{\pi(k-1)^2}^a(I), k \in \mathbb{N}^+ \quad (9)$$

Mohana and Rao showed, that area granulometry outperforms granulometry in accuracy. As we can see in figure 5b, there is a clear peak at $k = 21$. Also, there exists an efficient algorithm for area granulometry given by Meijster et al. [18, 19], that performs area-granulometry as a single opening.

However, the results are not always that clear. For instance the images figure 5c and 5e show fields of red blood cells in which the single cells are not clearly separated. This reflects directly in the result of area granulometry. The many overlapping cells in image 5c produce a slightly shifted peak in 5d, where as the structure of 5e produces two peaks in 5f.

3.2.3 Attribute and Shape-only Granulometry

Shape-only granulometries have been introduced by Urbach and Wilkinson [33], which is an extension of attribute granulometry as defined by Breen and Jones [5]. Urbach and Wilkinson define shape granulometry or shape spectrum as a spectrum generated using scale-invariant shape filters, which are thinnings (see section 3.1.3).

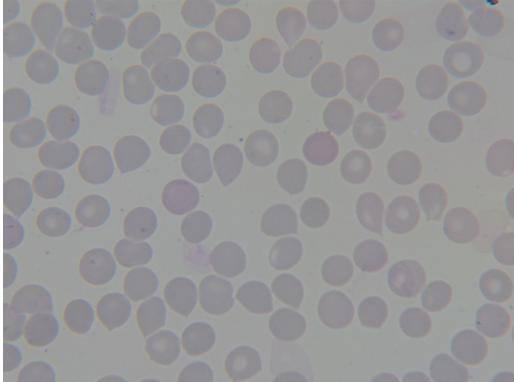
Increasing and not-increasing attributes (openings and thinnings, respectively) can be used to compute a generalized attribute granulometry, analogously to area granulometry in equation 9 [5]:

$$g^t(k) = \sum \gamma_{T_k}^t(I) - \sum \gamma_{T_{k-1}}^t(I), T_k \in A, k \in \mathbb{N}^+ \quad (10)$$

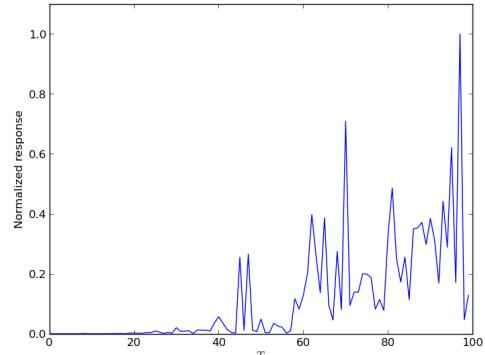
where $A = \{T_i, i = 1, 2, \dots\}$. A must be ordered in such a way that, if $C \subseteq M$ cannot satisfy any $T_i \in A$, it also cannot satisfy any $T_j \in A, j > i$ [5].

For thinnings, the granulometry would peak when the entire image is opened or closed. When this happens, the information provided by attribute granulometry would become meaningless. To counter this, we will impose an upper limit s_{max} to the size of connected components: if a connected component reaches a size greater than s_{max} , it will not grow any further. For our purpose, s_{max} is set to 20% of the image's total area.

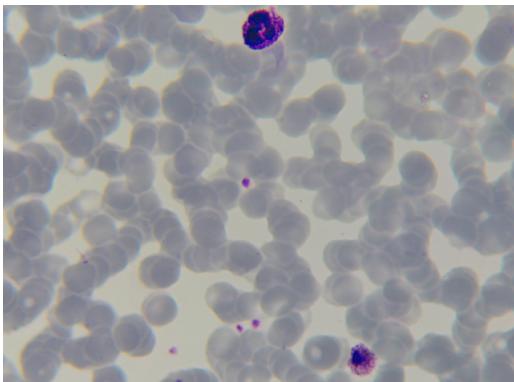
Let $B(C)$ denote the bounding box of a connected set C , $B(C)_w$ the width and $B(C)_h$ the height of the bounding box. We define two scale invariant



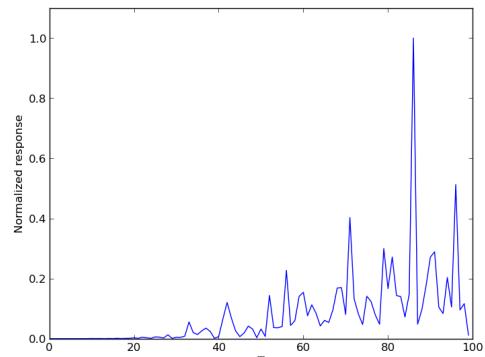
(a)



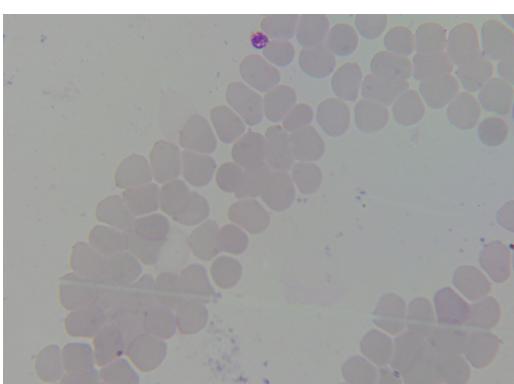
(b)



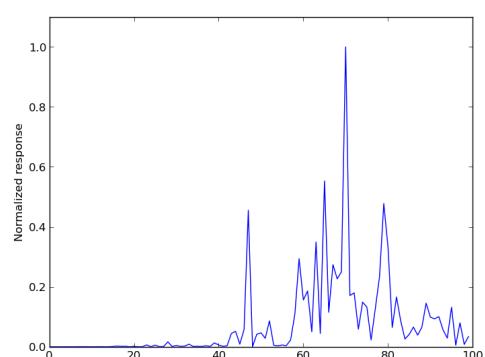
(c)



(d)



(e)



(f)

Figure 6: Normalized bounding box aspect ratio granulometries.

ordered sets of attributes, similar to Hernandez et al. [13]. The first one is aspect-ratio of the bounding box, A^r :

$$A^r = \{T_i \mid i = \frac{\min(B(C)_w, B(C)_h)}{\max(B(C)_w, B(C)_h)}\} \quad (11)$$

Here, i is a real number in $[0, 1]$, indicating the equality of both bounding box sides. The plotted aspect-ratio granulometries in figure 6 (projected onto $[0, 100]$) for our cell pictures show definite peaks for each image. Since most cells in figure 6a stand free, the global peak shown in figure 6b is close to 1 – disk-like components have square bounding boxes. The second highest peak around 0.7 indicates that there are elongated elements in the image, most probably cells that are connected.

In figure 6d aspect-ratio granulometry detects a high number of elements with a nearly square bounding box, even though the cells in figure 6c are overlapping heavily. This shift is comparable to the shift displayed in area granulometry (see figure 5c). Figure 6e indicates that the cells in the lower left corner of figure 6d are inseparable for the algorithm. Elements in the image are highly elongated.

The second ordered set of attributes which we define is the fill-ratio between actual area and bounding box, A^f . Let again $\|\cdot\|$ denote set cardinality.

$$A^f = \{T_i \mid i = \frac{\|C\|}{B(C)_w \cdot B(C)_h}\} \quad (12)$$

For a perfectly circular shaped connected component the expected ratio i_f is:

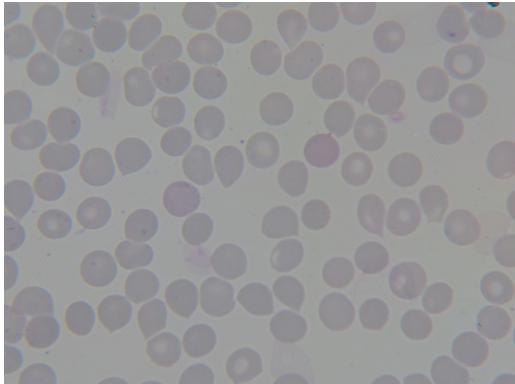
$$i_f = \frac{\pi r^2}{4r^2} = \frac{\pi}{4} \approx 0.79 \quad (13)$$

In figure 7b, we can see clearly that fill-ratio granulometry peaks around $i \approx i_f$. Still, there are many local peaks in the spectrum that indicate many elements for which this attribute computes a smaller ratio.

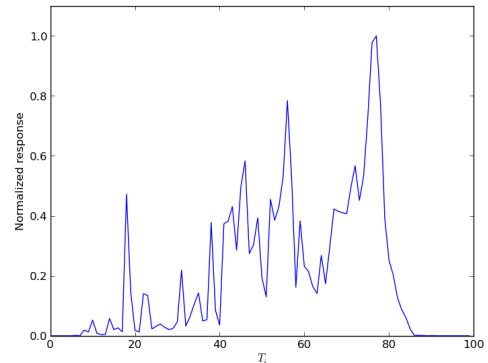
However, for images where cells are squashed or overlap, as in figure 7c and figure 7d, this granulometry computes peaks much lower than expected. This shows that fill-ratio opening will probably not produce any good results on these kinds of images.

4 Ultimate Attribute Opening of Red Blood Cell Images

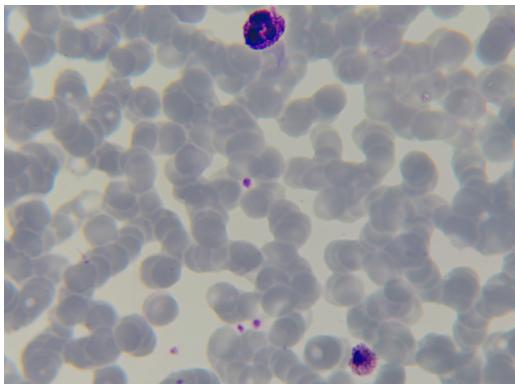
Ultimate Openings and Ultimate Attribute Openings, respectively, are parameter-free morphological operators on gray-scale images. This operator and its use for segmentation was first proposed by Beucher [3] and extended upon by Hernandez et al. [12, 13].



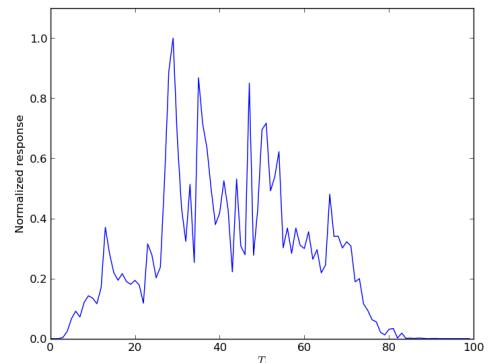
(a)



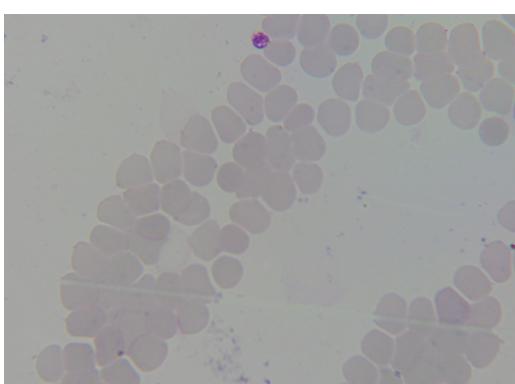
(b)



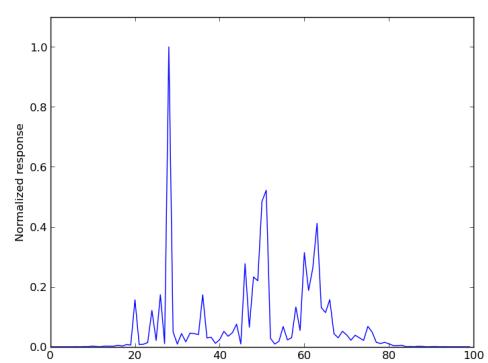
(c)



(d)



(e)



(f)

Figure 7: Normalized bounding box fill ratio granulometries.

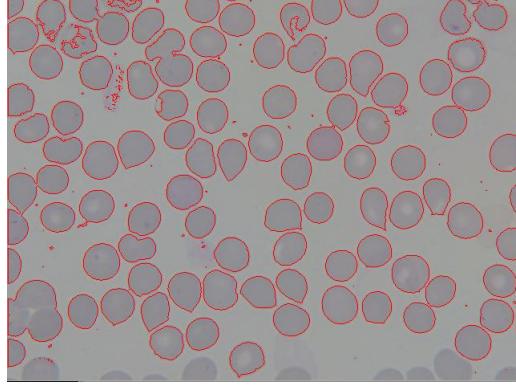


Figure 8: Naive foreground model.

The general idea of Ultimate Attribute Openings is to use the attribute opening of an image I which produces the greatest gray-scale residue.

We find this maximal residue by selecting λ to be the global maximum from the image's attribute granulometry. We then use λ to compute an attribute opening (or closing). The difference image between this ultimate opening to the original provides us with an optimized segmentation of the image.

In this section, we will take a look at segmentation of red blood cells using ultimate attribute openings, as an extension of ultimate openings as introduced by Beucher [3]. We will, however, not concern ourselves with shape ultimate attribute openings as proposed by Hernandez et al. [12, 13].

4.1 Ultimate Attribute Opening

Beucher introduced ultimate openings for standard morphological opening and closing, using a structuring element [3]. The idea is to perform a number of openings on an image I for a structuring element S of increasing size and returning the opening with the highest gray-scale residue, i.e. the opening producing the biggest change in the image. Essentially, this is computing the peak value of the granulometry of I [17].

We re-define ultimate attribute closing in our own notation. Let I be a gray scale image on domain M , A be an ordered set of attributes $T_i, i \in \mathbb{N}$ and \sum denote again the gray-scale sum of an image. Let moreover

$$\lambda = \operatorname{argmax}_k (\sum \varphi_{T_k}^t(I) - \sum \varphi_{T_{k-1}}^t(I)), T_k \in A, k \in \mathbb{N}^+ \quad (14)$$

the peak of the attribute granulometry of I (see section 3.2.3). Then the naive foreground model, i.e. the maximal residue of I , F' of I is:

$$F'(I) = I - \varphi_{T_\lambda}^t(I) \quad (15)$$

The $-$ operator denotes the pixel-wise difference between two images.

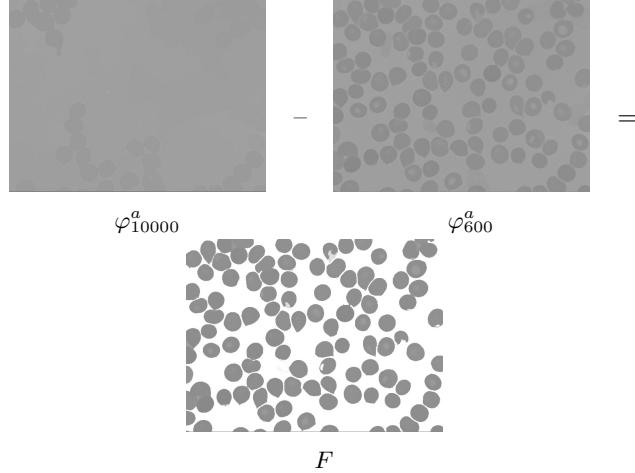


Figure 9: Step-wise computation of foreground model.

figure 8 shows such an naive foreground model. While $F'(I)$ succeeds to compute a foreground model that includes the red blood cells and also separates the majority of them correctly, it fails to remove small artifacts that are clearly not of any interest and also does not recognize the partially visible cells on the image border.

Let now moreover $\epsilon \in [0, \lambda[$. Then a better foreground segmentation algorithm can be defined by:

$$F(I) = \varphi_{T_\infty}^t(I) - \varphi_{T_{\lambda-\epsilon}}^t(I) \quad (16)$$

∞ denotes a pseudo-infinity: the only requirement for this number is that it is sufficiently greater than λ . We introduce ϵ to counter a natural variance in cell structure. Tek et al. [28] used a similar approach for area-based segmentation where they use $\epsilon = \frac{3\lambda}{4}$, stating the remaining size value is sufficient to find all cells present.

figure 9² illustrates the algorithm using area closing: the first closing computes a background model, where nearly all cells have been removed or their intensities have been lowered significantly. The second closing removes everything smaller in size than the cells. Their difference produces a foreground model of the scene.

To binarize $F(I)$ we use Otsu thresholding [22] and mark the cell borders using a sweep line algorithm to detect the external cell contours.

figure 10 shows the so computed foreground models for area closing, using $\epsilon = \frac{\lambda}{2}$ of three different cell images. In all images, there are no artifacts present in the foreground model. Figure 10a shows that the algorithm performs very

²For illustration purposes, the removed background, which actually is zero and colored black, is depicted as white.

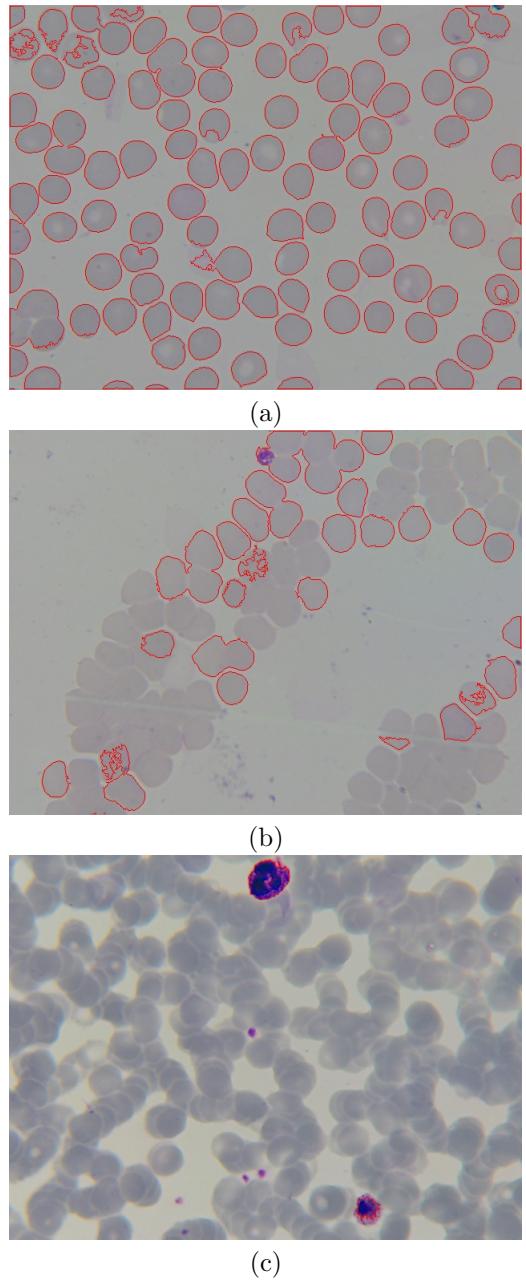


Figure 10: Segmentation for $\epsilon = \frac{\lambda}{2}$ using area closing.

good in finding the exact cell borders. This is due of the shape-adaptive nature of the dynamic structuring element that does not modify the morphology of the connected components in the image [35]. However, as shown in figure 10b and 10c, with degrading slide quality (i.e. how well a slide has been prepared), the algorithm's performance gets dramatically worse and it fails to extract the correct boundaries between touching or overlapping cells.

4.2 Ultimate Attribute Opening on RGB Color Channels

Until now, we used a gray-scale version of the color image by converting it from *RGB* to *HSV* color space and then selecting only the value component of it. This produces a gray-scale image that preserves most detail from the original.

While this is a good approach, it has the side effect of merging the information present in the separate color channels of the original *RGB* image. However, an interesting property of microscopy images (and generally all images taken through optical lenses) is, that light of different wave-lengths, i.e. the colors, focus in different places: in front, directly on and behind the charge-coupled device (*CCD*). This effect is called chromatic aberration [10]. Modern lenses use filters to counter this effect, but due to the magnification of microscopes, chromatic aberration becomes visible when using *CCD* cameras.

Other studies, as for example performed by Zoueu et al. [39], also use physical and representational color image properties to analyze Malaria slides without requiring staining.

Our images are stained using Giesma stain [38], which binds to certain proteins and colors them violet, as described in section 1.1. We will in this section look at if we can exploit the resulting image properties to improve the segmentation algorithm proposed in section 4.1.

The approach is simple: analyze each color channel component of the *RGB* image separately! Some color channels convey more useful information than others, which are shadowed by the value channel of the image in *HSV* color space.

figure 11 shows the four color channels in question as gray-scale images. Figure 11a represents the value component of the *HSV* image, while 11b through 11d represent the *RGB* components, ordered blue, green and red.

In figure 11b we can see that the image is not entirely focused. Cells, that are clearly separated in image figure 11a are fading into each other slightly. The green channel 11c exhibits the highest resolution of all channels. It does, however, introduce some artifacts that are not or to a lesser degree present in the other channels. The red channel 11d is also slightly out of focus, but to a degree that produces a greater separation between the single cells. At the same time, the artifacts found in 11c are only visible to a negligible degree.

These qualitative differences between the single channels of the *RGB* color space have an impact on the segmentation performance of ultimate attribute closing.

In figure 12 we can see how ultimate attribute closing performs on the single *RGB* channels. While the results in images figure 12a through 12c are similar,

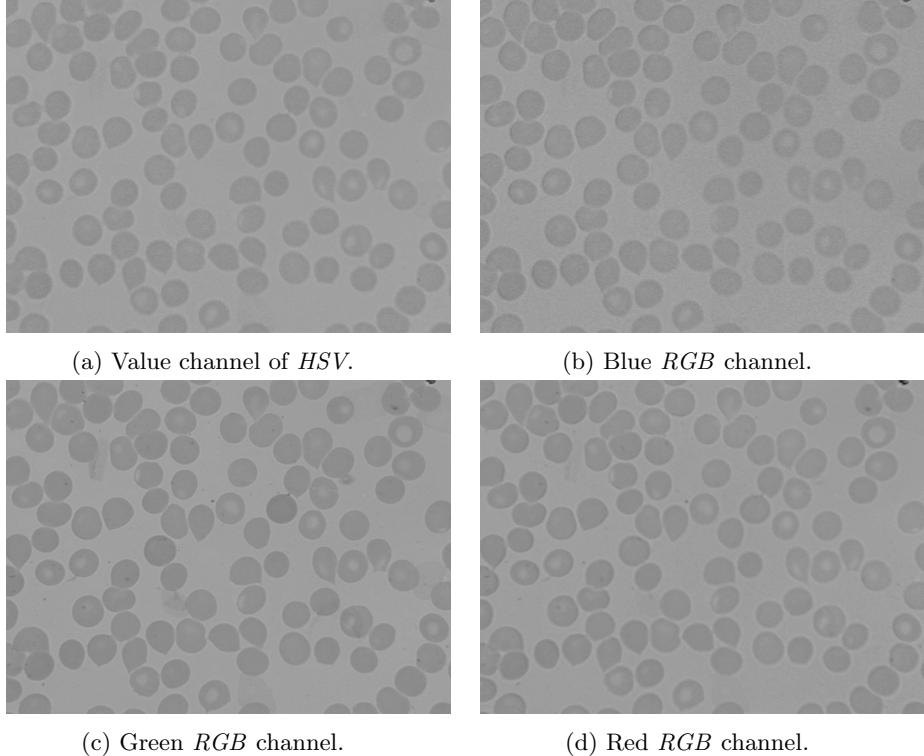


Figure 11: Separate color channels of the same image.

the red color channel clearly performs best. Still, its performance is still very similar to the original segmentation in figure 10a. It also fails to separate groups of cells that are touching, as visible in the lower left corner of the image.

This is different for images figure 12d through 12f. Here, the original segmentation (figure 10b) performs closely to figure 12e, while 12d and 12f produce clearly a better segmentation segmentation result – more cells are extracted from the compound cell areas in the upper right and lower left corner of the image. These foreground models however also include some elements that have only been partially extracted.

Also images with overlapping cells, as displayed in figure 13, benefit from segmenting them only on the blue and red color channels (13a and 13c). Nevertheless, the produced foreground models are still highly fragmented and certainly of no proper use.

5 Implementation

In this section, we will take a look at some implementation details of the used algorithms. This report contributes by providing an efficient open-source im-

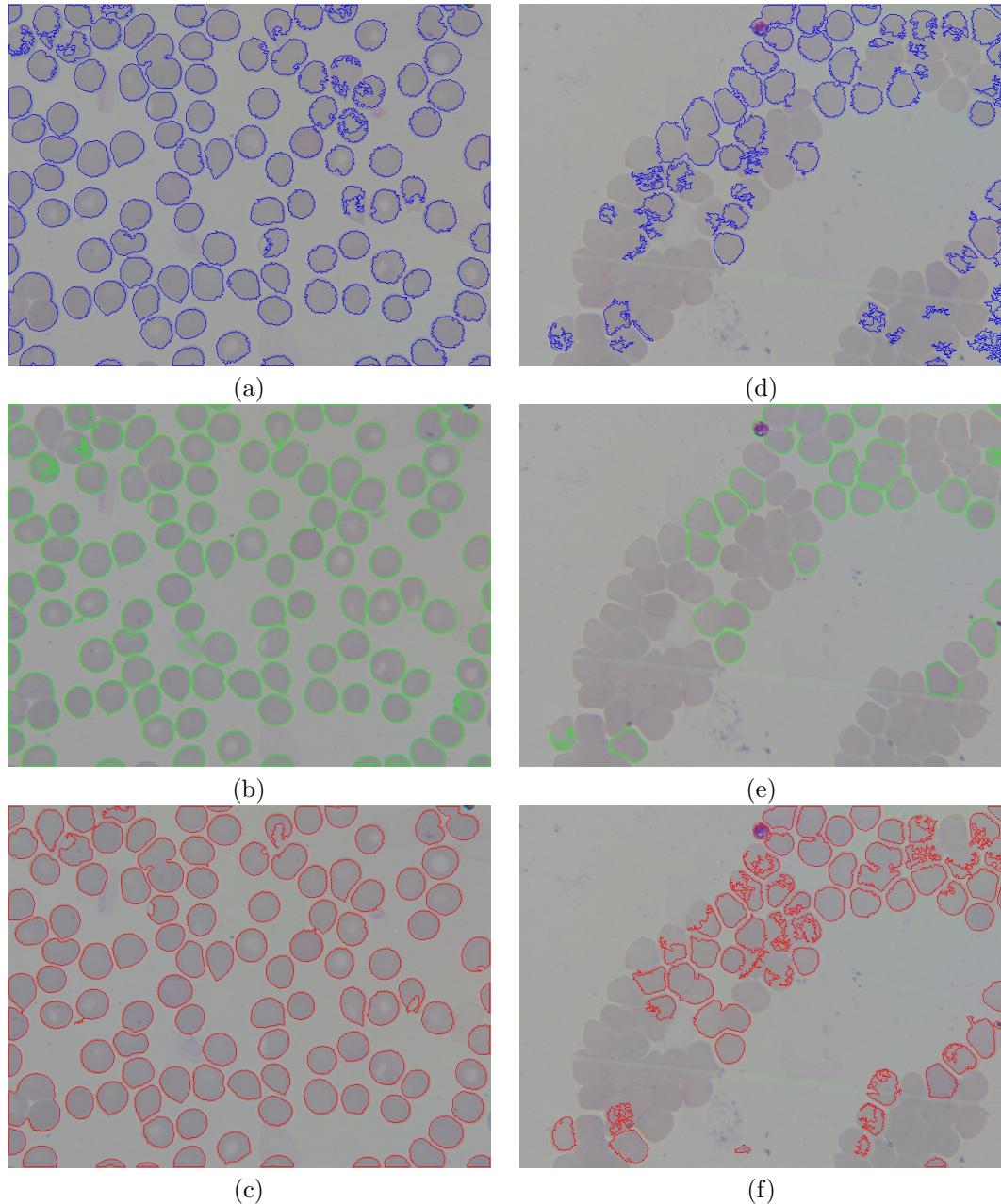


Figure 12: Separate color channel segmentation for $\epsilon = \frac{\lambda}{4}$ using area closing.
The resulting masks are filtered with an area closing of size $\lambda = 150$.

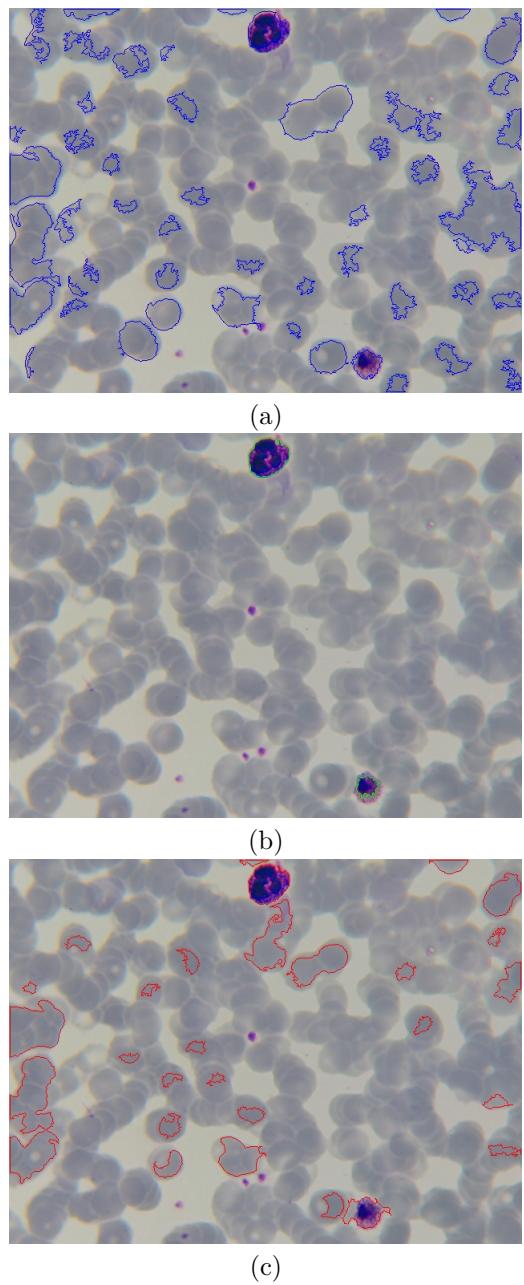


Figure 13: (Continued from figure 12). Separate color channel segmentation for $\epsilon = \frac{\lambda}{4}$ using area closing. The resulting masks are filtered with an area closing of size $\lambda = 150$.

lementation of morphological attribute opening [5] as proposed by Meijster et al. [19, 37] on gray-scale images. This implementation has also been extended to perform attribute granulometry and thinning [18, 19]. The implementation is done in C++ and is compatible with the popular OpenCV³ libraries. It can be downloaded at <http://bitbucket.org/fbie/morphology>.

Morphological attribute opening and closing are algorithms that are not provided by the OpenCV libraries. There exists an implementation for Matlab, which only accepts binary images as input⁴. Other implementations are not available for free, as they are part of libraries that are only commercially available⁵. These implementations are based on Max-Trees instead of union-find.

All other operations used in this report, like morphological operations with structuring elements, thresholding, logical or arithmetic operations on images or contour detection on binary images, have been provided by OpenCV.

5.1 Attribute Opening and Closing

Mejster et al. and Wilkinson et al. proposed an algorithm to compute area openings and closings efficiently, using a union-find data structure [9, 19, 37]. Union-find represents disjoint sets using trees and is used in computer vision to find connected elements in binary images [26, 4].

The key to union-find is to compress the path from a leaf-node to its root node in the union-find tree as soon as possible, which amortizes the complexity of finding the root to $O(6)$ [9].

Area opening and closing works through elements of the gray-scale image in order of their respective gray-values. The difference between both is that the order in which they process the pixels is reversed [19].

Their algorithm first sorts all pixels in the image after in- or decreasing value and level-pixels after scan-line order (i.e. from the upper left corner to the lower right corner of the image) and turns every pixel into a mathematical singleton. Then, the list is iterated: we look at each pixel set and each of its 8-connected neighbors, as long as they are sorted before the currently selected pixel. Level pixels, i.e. pixels of the same gray value, are merged directly: their corresponding sets are united. Non-level pixels are merged conditionally with their neighbors only if the area of the set they belong to is less than λ [19, 37].

The union routine used is asymmetrical, to ensure that the last visited pixel, i.e. the one with the highest or lowest gray value respectively, is the root of the set representation. This enables us to resolve the gray-scale values by iterating reversely over the pixels in the sorted list, assigning each pixel the value of its parent. Since all sets are smaller or equal to λ in size, the connected components are assigned the value of their surrounding component [19, 37].

Hereby, the structure of the connected components of size greater than λ is preserved, as they are simply not represented in the disjoint set structure yet.

³<http://www.opencv.org>

⁴<http://www.mathworks.se/help/images/ref/bwareaopen.html>

⁵<http://www.mmorph.com/cppmorph/>

This algorithm can easily be extended to attribute opening and closing, where each pixel set is associated with an attribute [19].

The implementation used for this report does not rely on arrays, as originally proposed. Instead, a pointer-structure represents the relations between pixels.

5.2 Attribute Granulometry

Meijster et al. extended their algorithm to computer granulometries in much the same fashion as it computes openings and closings [19, 18].

In this version, the algorithm adds the gray-level difference between the current and the new root pixel to the spectrum for the according attribute when two sets are united. The final step to resolve gray-scale values is omitted and instead of an image the algorithms returns the differential spectrum of gray-values over attributes. This means that attribute granulometry can be implemented efficiently as a single opening instead of naively from its definition [19, 18]. The definition of attribute granulometry suggests that otherwise, we would have to compute as many openings as there are attributes until the openings converge, or an upper limit of attributes is reached.

Meijster et al. do not directly point out that their algorithm for granulometries is applicable to attributes, but this can be easily inferred from the code examples they provide. The only requirement is that the attributes must be represented as a scalar.

6 Discussion

In this section, we will discuss the segmentation performance of the proposed method. Additionally, we will discuss how the performance could be improved and what can be identified as future work.

6.1 Analysis of Segmentation Performance

In section 4.2, we already saw a few examples for the ultimate attribute closing on *RGB* color channels for separated, squashed and overlapping blood cells.

As outlined in section 1.2, we frequently encounter four types of images when analyzing blood sample images. These show cells that are:

- a) Separated
- b) Squashed
- c) Overlapping
- d) Out of focus

These conditions are due to human factors during sample preparation and image capturing. Squashed and overlapping cells are a result of bad preparation of the samples [38]. It is, however, hard to prepare samples correctly and it

requires training. These features are permanent as soon as the sample has been fixed on the glass slide.

Opposing this, wrong focus on cells is a matter of image capturing. The focus can still be corrected after preparing the samples and also not as hard a task given the necessary equipment.

In section A in the appendix of this report, there are twelve sample images and the performance of the proposed algorithm showcased. Each image is processed using the foreground segmentation method as proposed in section 4.2 using area as an attribute and the attributes proposed in section 3.2.3. Each image is segmented once for each color channel and for each attribute. This results in twelve segmentations per image.

Separated Blood Cells Images of well separated cells that are properly focused are obviously the most easy to handle cases. The first three image samples show images of well separated cells. Here, we can see how the attribute closings behave differently. Area closing tends to separate regions and sometimes results in over-segmentation. Still, other parts of the image, like tightly touching cells, are under-segmented.

Aspect ratio (equation 11) and fill ratio (equation 12) attributes tend to group elements and therefore tend to produce under-segmentations. For example, the image used in figure 14 shows a dense structure in the lower right corner which is inhabited by parasites (violet features). Using aspect-ratio and fill-ratio, these cells are always merged into one region. In the red color channel, however, area closing divides those into separate regions. Still, area closing fails to detect the entire cells.

In figure 15 we can see how aspect-ratio and fill-ratio closings fail to detect any cells at all on the image. This is true for the value component of the *HSV* representation of the image as well as for the red color channel. This illustrates how much the color channels differ from each other and how the coloring of images taken from the same sample can differ over areas (due to variances of staining and background illumination) [27].

The last sample image, figure 16, again shows how area closing over- and bounding-box based closings under-segment blood cells.

From all three examples we can see that artifacts are not posing a problem in segmentation. They are either segmented separately and can be therefore classified as random artifacts later on, or inside of cells, in which case the segmentation behaves correctly.

Generally seen, aspect-ratio and fill-ratio closings result in nearly identical segmentations. As their differences are so small, we will refer to them as bounding-box based closings in the subsequent text.

Overlapping Blood Cells Overlapping blood cells pose a major difficulty for the proposed segmentation. The algorithm requires clear cell borders to separate the single cells. As already shown in figure 5d, 6d and 7d, the granulometries for

images with overlapping cells peak at points differing from those for optimally separated cells.

The segmentations resulting from these granulometries (figure 17, 18 and 19) show clearly that the algorithm generates mainly under segmented regions. Additionally, area closing produces over-segmentations. Nevertheless, bounding-box based closings for the red color channel separate fore- and background correctly. In all three examples, the segmentation on the red color channel performs as least as good as the other three color channels.

Squashed Blood Cells As with overlapping blood cells, squashed blood cells do not exhibit strong borders. Instead, their borders touch with those of other cells, building compound structures. The only indication of multiple cells is the outer curvature which shows concave peaks.

However, we can circumvent these difficulties by taking advantage of chromatic aberration. By selecting a color channel which is focused in such a way that the single blood cells get smaller, we increase the distances between cells and therefore the cells become better separated.

In figure 20, 21 and 22 we see how this works in practice. In figure 20 again the tendency of bounding-box based closings of under-segmenting images is exhibited. However, neither of the segmentations manages to separate the compound red blood cells in the upper center of the sample image.

Figure 21 shows that, in some cases, area closing on the red color channel can outperform segmentations on other color channels or using different closing types. The *HSV* value channel as well as the blue and green *RGB* channel do not respond accurately to area closing. Bounding-box based segmentations result again in under-segmentation. Note that here, for the first time, artifacts play a role in the segmentation. In the blue color channel, all attribute closings select a thin elongated element (central left part of the image) as part of the foreground, while clearly, it is not a red blood cell.

Nevertheless, in figure 22 we can see, again, that area closing on the red color channel can fail to detect compound cells at all. This is due to the area of these compounds being larger than our pseudo-infinity, which is used to compute the background model as illustrated in figure 9. Therefore, the structure is being removed entirely when computing the pixel-wise difference. Here, bounding-box based segmentations produce better results, even though they produce under segmentations.

Blurred Blood Cells Segmenting blurred cells is somewhat of a corner case, as the sharpness of images is a factor that is easy to influence during image acquisition if given the right tools. However, for completeness and because of some setups, where this actually might occur (e.g. image capturing in rapid succession or cells residing in different layers of the blood on the glass slide), we will here also analyze those images.

The examples of blurred blood cells in figure 23 shows how much the blurring influences the segmentation performance. Except for the blue color channel, no

other channel provides enough information for a proper segmentation. This is mostly the case because the histogram is not broad enough for Otsu thresholding. The performance of bounding-box closings is somewhat acceptable if we only want to find a foreground model. Still, single cells can not properly be extracted.

In contrast, figure 24 showcases that cells with blurred boundaries can very well be extracted, if the contrast between background and cells is still high enough. The red color channel segmentation using area closing shows, that cells that have intensities close to the background color are still not extracted correctly. The overall performance is comparable to the segmentations of focused separated cells.

Overall good red color channel performance can again be observed in figure 25. However, multiple layers of cells (not overlapping) and as a result different focus and intensity levels of cells make it hard for area granulometry to perform correctly – therefore, area closing performs worst of all red color channel segmentations. Again, bounding-box closings result in under-segmented regions.

6.2 Future Work

Summing up the observations made in section 6.1, there are two things that stick out:

- The red color channel of cell images in *RGB* color space performs most reliably.
- Bounding-box based attribute closings tend to under-segment images.

However, variances in coloring can make it impossible for the algorithm to perform reliably on the red color channel. The latter point hints that bounding-box based segmentation yields a reliable foreground-background segmentation.

Images, where cells are well separated and where the camera was focused correctly can be segmented most reliably. The value channel of *HSV* and the green color channel of *RGB* images perform similarly.

Based on these observations, we will now take a look at future work. There are two different types of future work relevant for what has been presented in this report: improvements on the algorithm’s performance and the continuation of evaluating results.

6.2.1 Improvements

While the method we saw in section 4.1 and 4.2 produces generally good results on well prepared blood slides, there are issues. Simply touching cells are sometimes not separated correctly. Also, for images of different qualities, as discussed in section 6.1, the performance is not constant.

Under-segmentation is one of the main issues. Ultimate attribute opening is a powerful operator, but it tends to shadow relevant structures by their super sets [13]. Hernandez et al. introduced shape ultimate attribute opening, which

increases the value of residues if its elements match a given shape description. They introduce shape ultimate attribute openings as a general segmentation tool but also propose its use for cell segmentation. While they state that slightly touching cells are not always separated by their technique[13], it could help segmenting bigger clusters of cells. An improved version of the algorithm presented here could also take shape parameters into account.

Moon et al. use the curvature of cells to detect if an element is a compound cell [21]. An improved ultimate attribute closing segmentation algorithm could also use curvature as an attribute. However, maintaining information about curvature about each connected component for union-find based attribute opening is computationally expensive. Polygon-union algorithms and approximations could be a great asset in this case.

Regarding color variances, Tek at al. used a normalization method for *RGB* images of malarial blood samples [27]. Applied to color channel based segmentation like the one presented here, this could prevent the algorithm failing if there is not enough red color present in the image.

As seen in section 6.1, bounding-box based attributes are good at producing general foreground segmentations. Therefore, a linear combination of bounding-box segmentation followed by area closing could be used to optimize the result. Combining results of the segmentation on the blue and red color channels could also produce good results.

These suggestions for improvement are identified as future work.

6.2.2 Evaluations

This report does not cover a quantitative evaluation of the method proposed. This is due to two major factors: availability of working implementations of methods from literature and data availability. Both of these points need to be addressed in order to perform a proper and unbiased evaluation.

Methods need to be re-implemented. Working implementations of methods from literature are mostly not directly available for evaluation. Implementing them from literature would consume quite some time, which is beyond the scope of this report. However, this would be necessary for properly comparing the performance of cell segmentation algorithms. This is an important task for future work on this topic.

Comparing the algorithm presented in this report to standard methods would be possible. Still, such a comparison would mainly serve to lure the reader into thinking that the proposed method is a superior approach, without presenting proper evidence, and is therefore unfeasible.

Large amounts of manually annotated data is necessary. For an unbiased evaluation, a collection of annotated images, which contains number of cells present, condition of the image and segmentation ground truth would be required.

Hernandez et al. point out that there is no library of annotated images that could be used for a quantitative evaluation. They add that the creation of such a library is out of scope for their work [13], which is also true for this project. The domain of Hernandez et al. is, however, different from ours.

In his dissertation, Tek presents a manual multiple thresholding selection method and evaluation algorithm for generating segmentation ground-truth for single images [28]. However, he does not annotate the conditions present in the images. Most publications in general waive to elaborate on the quality of the images used to evaluate algorithms.

7 Conclusion

In this report, we witnessed the construction and qualitative analysis of an ultimate attribute opening [3, 13] and color channel based algorithm for segmenting cells in peripheral Malaria blood slides. The algorithm computes the optimal attribute with which the image should be filtered and is run on different *RGB* color channels.

After analyzing the problem, we reviewed associated literature on automated Malaria diagnosis and segmentation of blood cells in microscopy images. We concluded that there are many different approaches, where morphology plays a dominant role, but that there is no widely accepted solution to the problem.

We saw a detailed description of the foundations for the method used, namely mathematical morphology. We looked at the properties of morphological openings and closings with structuring elements, openings and closings using dynamic structuring elements [35] and their generalizations, attribute openings and closings [5]. We also learned about tools associated with these filters, namely attribute granulometry [5], which provide us with general information about the structure of a gray-scale image. This information can be very valuable in the further processing of gray-scale images.

Furthermore, this report featured a formalized description of the morphological segmentation method. Difficulties and possible solutions have been reflected upon and segmentation results have been presented. Images of slides that have been prepared well are easy to analyze using this method, but difficulties arise when the slide quality degrades.

Additionally, we learned about the implementation details of the algorithms. The implementation used bases on the work of Meijster et al. [18, 19]. As a side effect of this project, an open source implementation of morphological attribute opening is now available and compatible with OpenCV.

Finally, we discussed the nature of the results conceived during experimentation and saw a number of directions available when considering future work on this topic.

As a conclusion, we can see that color channel based ultimate attribute opening can be a good segmentation tool for images of red blood cells. If the conditions under which the image has been acquired are less controllable, it can not provide a stable segmentation. However, if extended to favor certain shapes

by using shape ultimate attribute closings [13] the performance could possibly be improved. Pre-processing the images with color-normalizing methods could also be a good measure for stabilizing segmentation performance [27].

Segmentation of blood cells for Malaria analysis is an important step towards faster and more accessible diagnosis of patients. Automated Malaria analysis can reduce the time researchers are required to perform tedious manual laboratory tasks. The ultimate goal, however, would be a system for Malaria diagnosis that can be accessed remotely even by non-experts. Such a system would be a major step towards fighting Malaria.

References

- [1] D. Anggraini, A. S. Nugroho, C. Pratama, I. E. Rozi, A. A. Iskandar, and R. N. Hartono. Automated status identification of microscopic images obtained from malaria thin blood smears. In *Electrical Engineering and Informatics (ICEEI), 2011 International Conference on*, pages 1–6. IEEE, July 2011.
- [2] T. Bergen, D. Steckhan, T. Wittenberg, and Thorsten Zerfass. Segmentation of leukocytes and erythrocytes in blood smear images. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 3075–3078. IEEE, August 2008.
- [3] Serge Beucher. Numerical residues. *Image Vision Comput.*, 25(4):405–415, April 2007.
- [4] Gary R. Bradski. *Learning OpenCV : computer vision with the OpenCV library*. O'Reilly, 2008.
- [5] Edmond J. Breen and Ronald Jones. Attribute Openings, Thinnings, and Granulometries. *Computer Vision and Image Understanding*, 64(3):377–389, November 1996.
- [6] Cecilia Di Ruberto, Andrew Dempster, S. Khan, and B. Jarra. Segmentation of blood images using morphological operators. In *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, pages 397–400. IEEE Comput. Soc, 2000.
- [7] Cecilia Di Ruberto, Andrew Dempster, Shahid Khan, and Bill Jarra. Morphological Image Processing for Evaluating Malaria Disease. In Carlo Arcelli, LuigiP Cordella, and GabriellaSamiti Baja, editors, *Visual Form 2001*, volume 2059 of *Lecture Notes in Computer Science*, pages 739–748. Springer Berlin Heidelberg, 2001.
- [8] Cecilia Di Ruberto, Andrew Dempster, Shahid Khan, and Bill Jarra. Analysis of infected blood cell images using morphological operators. *Image and Vision Computing*, 20(2):133–146, February 2002.

- [9] Harold N. Gabow and Robert E. Tarjan. A linear-time algorithm for a special case of disjoint set union. *Journal of Computer and System Sciences*, 30(2):209–221, April 1985.
- [10] Robert D. Guenther. *Modern optics*. Wiley, January 1990.
- [11] S. Halim, T. Bretschneider, Y. Li, P. R. Preiser, and C. Kuss. Estimating Malaria Parasitaemia from Blood Smear Images. In *Control, Automation, Robotics and Vision, 2006. ICARCV 2006. 9th International Conference on*, pages 1–6. IEEE, December 2006.
- [12] Jorge Hernández and Beatriz Marcotegui. Ultimate Attribute Opening Segmentation with Shape Information. In MichaelH Wilkinson and JosB Roerdink, editors, *Mathematical Morphology and Its Application to Signal and Image Processing*, volume 5720 of *Lecture Notes in Computer Science*, pages 205–214. Springer Berlin Heidelberg, 2009.
- [13] Jorge Hernández and Beatriz Marcotegui. Shape ultimate attribute opening. *Image and Vision Computing*, 29(8):533–545, July 2011.
- [14] V. V. Makkapati and R. M. Rao. Segmentation of malaria parasites in peripheral blood smear images. In *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, pages 1361–1364. IEEE, April 2009.
- [15] M. T. Makler, C. J. Palmer, and A. L. Ager. A review of practical techniques for the diagnosis of malaria. *Annals of Tropical Medicine and Parasitology*, pages 419–433, June 1998.
- [16] S. Mandal, A. Kumar, J. Chatterjee, M. Manjunatha, and A. K. Ray. Segmentation of blood smear images using normalized cuts for detection of malarial parasites. In *India Conference (INDICON), 2010 Annual IEEE*, pages 1–4. IEEE, December 2010.
- [17] P. Maragos. Pattern spectrum and multiscale shape representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):701–716, July 1989.
- [18] A. Meijster and M. H. F. Wilkinson. Fast computation of morphological area pattern spectra. In *Image Processing, 2001. Proceedings. 2001 International Conference on*, volume 3, pages 668–671 vol.3. IEEE, 2001.
- [19] A. Meijster and M. H. F. Wilkinson. A comparison of algorithms for connected set openings and closings. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(4):484–494, April 2002.
- [20] K. N. R. Mohana Rao and A. G. Dempster. Area-granulometry: an improved estimator of size distribution of image objects. *Electronics Letters*, 37(15):950+, 2001.

- [21] Seunghyun Moon, Sukjun Lee, Heechang Kim, Lucio H. Freitas-Junior, Myungjoo Kang, Lawrence Ayong, and Michael A. E. Hansen. An Image Analysis Algorithm for Malaria Parasite Stage Classification and Viability Quantification. *PLoS ONE*, 8(4):e61812+, April 2013.
- [22] Nobuyuki Otsu. A Threshold Selection Method from Gray-Level Histograms. *Systems, Man and Cybernetics, IEEE Transactions on*, 9(1):62–66, January 1979.
- [23] Nicola Ritter and James Cooper. Segmentation and border identification of cells in images of peripheral blood smear slides. In *Proceedings of the thirtieth Australasian conference on Computer science - Volume 62*, ACSC '07, pages 161–169, Darlinghurst, Australia, Australia, 2007. Australian Computer Society, Inc.
- [24] Nicholas E. Ross, Charles J. Pritchard, David M. Rubin, and Adriano G. Dusé. Automated image processing method for the diagnosis and classification of malaria on thin blood smears. 44(5):427–436, 2006.
- [25] Jean Serra and Luc Vincent. An overview of morphological filtering. *Circuits, Systems, and Signal Processing*, 11(1):47–108, March 1992.
- [26] Richard Szeliski. *Computer Vision*. Springer London, London, 2011.
- [27] F. B. Tek, A. G. Dempster, and I. Kale. A Colour Normalization Method for Giemsa-Stained Blood Cell Images. In *Signal Processing and Communications Applications, 2006 IEEE 14th*, pages 1–4. IEEE, April 2006.
- [28] F. Boray Tek. *Computerised Diagnosis of Malaria*. PhD thesis, University of Westminster, September 2007.
- [29] F. Boray Tek, Andrew Dempster, and Izzet Kale. Computer vision for microscopy diagnosis of malaria. *Malaria Journal*, 8(1):153+, 2009.
- [30] F. Boray Tek, Andrew G. Dempster, and İzzet Kale. Parasite detection and identification for automated thin blood film malaria diagnosis. *Computer Vision and Image Understanding*, 114(1):21–32, January 2010.
- [31] Andrej Trampuz, Matjaz Jereb, Igor Muzlovic, and Rajesh M. Prabhu. Clinical review: Severe malaria. *Critical care (London, England)*, 7(4):315–323, August 2003.
- [32] E. R. Urbach, J. B. T. M. Roerdink, and M. H. F. Wilkinson. Connected Shape-Size Pattern Spectra for Rotation and Scale-Invariant Classification of Gray-Scale Images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(2):272–285, February 2007.
- [33] E. R. Urbach and M. H. F. Wilkinson. Shape-Only Granulometries and Grey-Scale Shape Filters. In Hugues Talbot and Richard Beare, editors, *Proceedings of the VIth International Symposium - ISMM 2002*, volume 6. CSIRO Publishing, April 2002.

- [34] Luc Vincent. Morphological grayscale reconstruction in image analysis: applications and efficient algorithms. *Image Processing, IEEE Transactions on*, 2(2):176–201, April 1993.
- [35] Luc Vincent. Morphological Area Openings and Closings for Grey-scale Images. In Ying-Lie, Alexander Toet, David Foster, HenkJ Heijmans, and Peter Meer, editors, *Shape in Picture*, volume 126 of *NATO ASI Series*, pages 197–208. Springer Berlin Heidelberg, 1994.
- [36] D. Wermser, G. Haussmann, and C. E. Liedtke. Segmentation of blood smears by hierarchical thresholding. *Computer Vision, Graphics, and Image Processing*, 25(2):151–168, February 1984.
- [37] Michael H. F. Wilkinson and Jos B. T. M. Roerdink. Fast Morphological Attribute Operations Using Tarjan’s Union-Find Algorithm. In *In Proceedings of the ISMM2000*, pages 311–320, 2000.
- [38] World Health Organization. *Basic Malaria Microscopy: Part 1 Learner’s guide*. World Health Organization, 20 Avenue Appia, 1211 Geneva 27, second edition, 1991.
- [39] Jeremie T. Zoueu, Georges L. Loum, T. Cisse Haba, Mikkel Brydegaard, and Herve Menan. Optical Microscope Based on Multispectral Imaging Applied to Plasmodium Diagnosis. *Journal of Applied Sciences*, 8(15):2711–2717, December 2008.

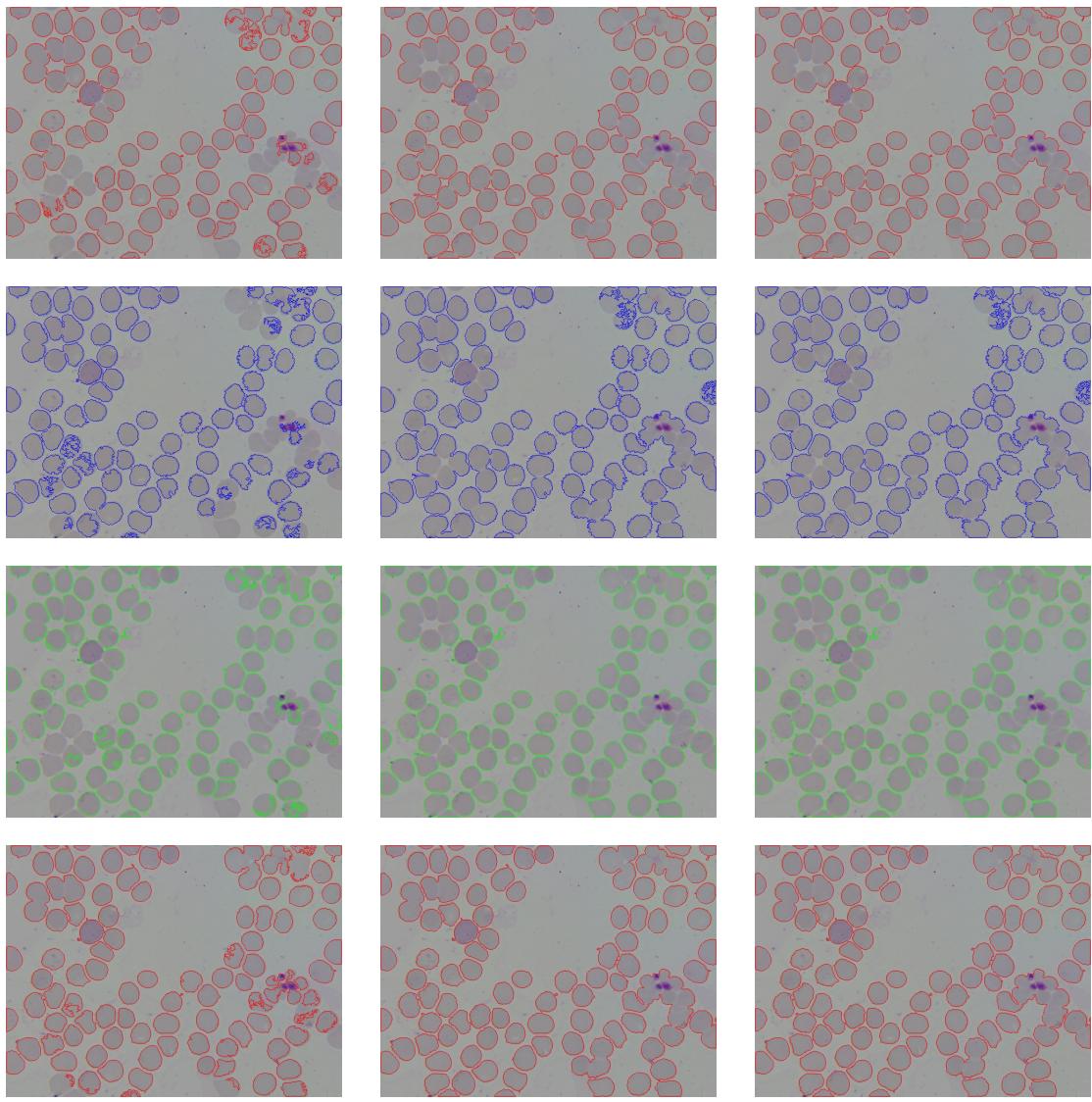
A Segmentation Results

This section contains segmentation results encoded as images from twelve different images of the image sets compiled by F. Boray Tek. The twelve images come from four different groups, as described in section 1.2, namely separated, overlapping, squashed and blurred cells.

Each image is presented as a matrix of images. The rows are ordered after color channels, namely value channel of the *HSV* color space, followed by the blue, green and red color channel of the *RGB* color space.

The columns depict the different attributes used for segmentation. All images have been segmented using the foreground model computation algorithm as presented in equation 16 in section 4.1 with $\epsilon = \frac{3\lambda}{4}$. The resulting segmentation images have been binarized using Otsu’s method and filtered with an area closing of size 150. The borders have been detected using a sweep-line algorithm and painted in the according colors on top of the original image.

If you do not have a colorized version of this report, you can obtain one at http://itu.dk/~fbie/morphological_segmentation_malaria.pdf.

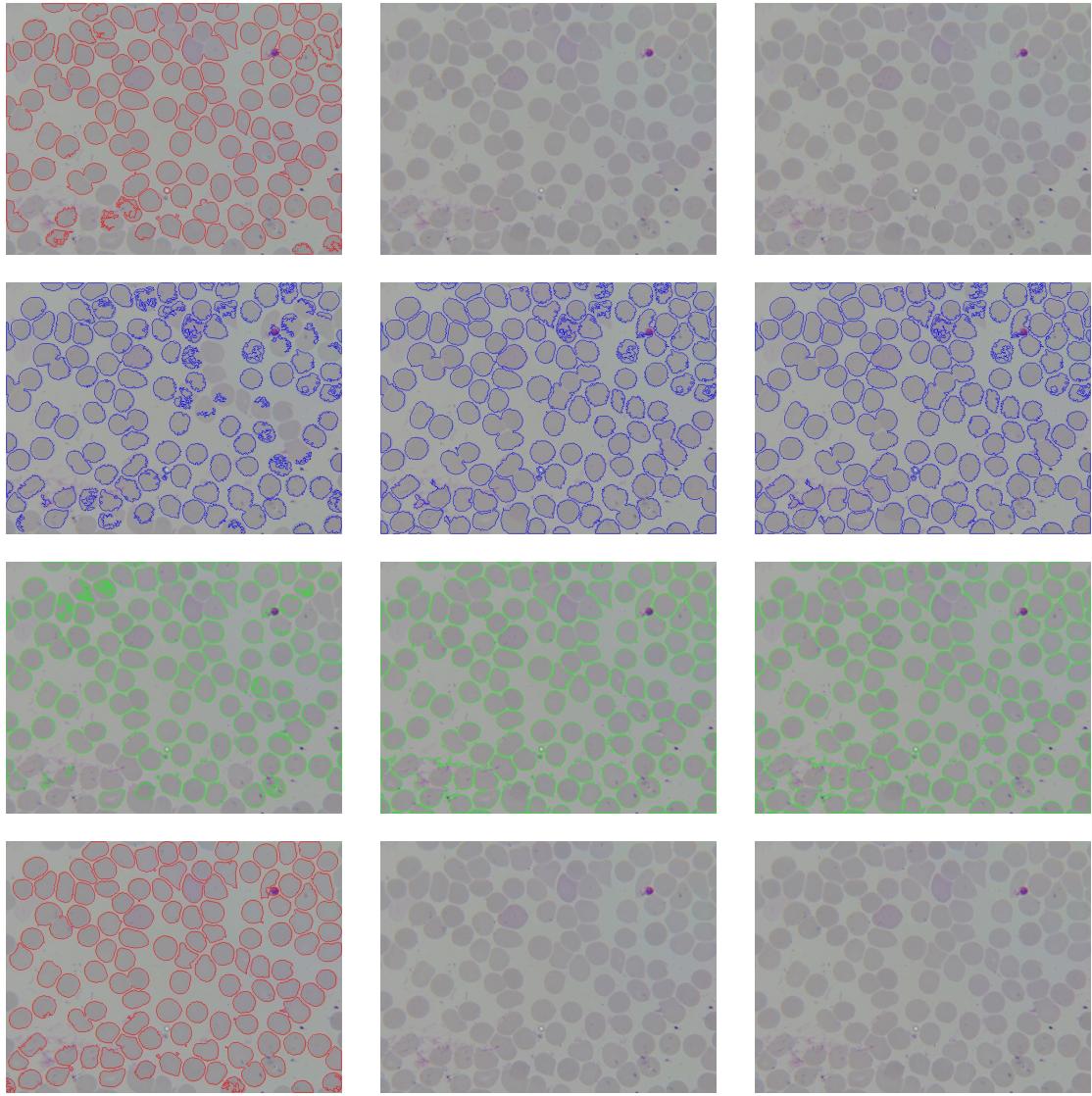


Area

Aspect ratio

Fill ratio

Figure 14: Segmentation results for separated cells

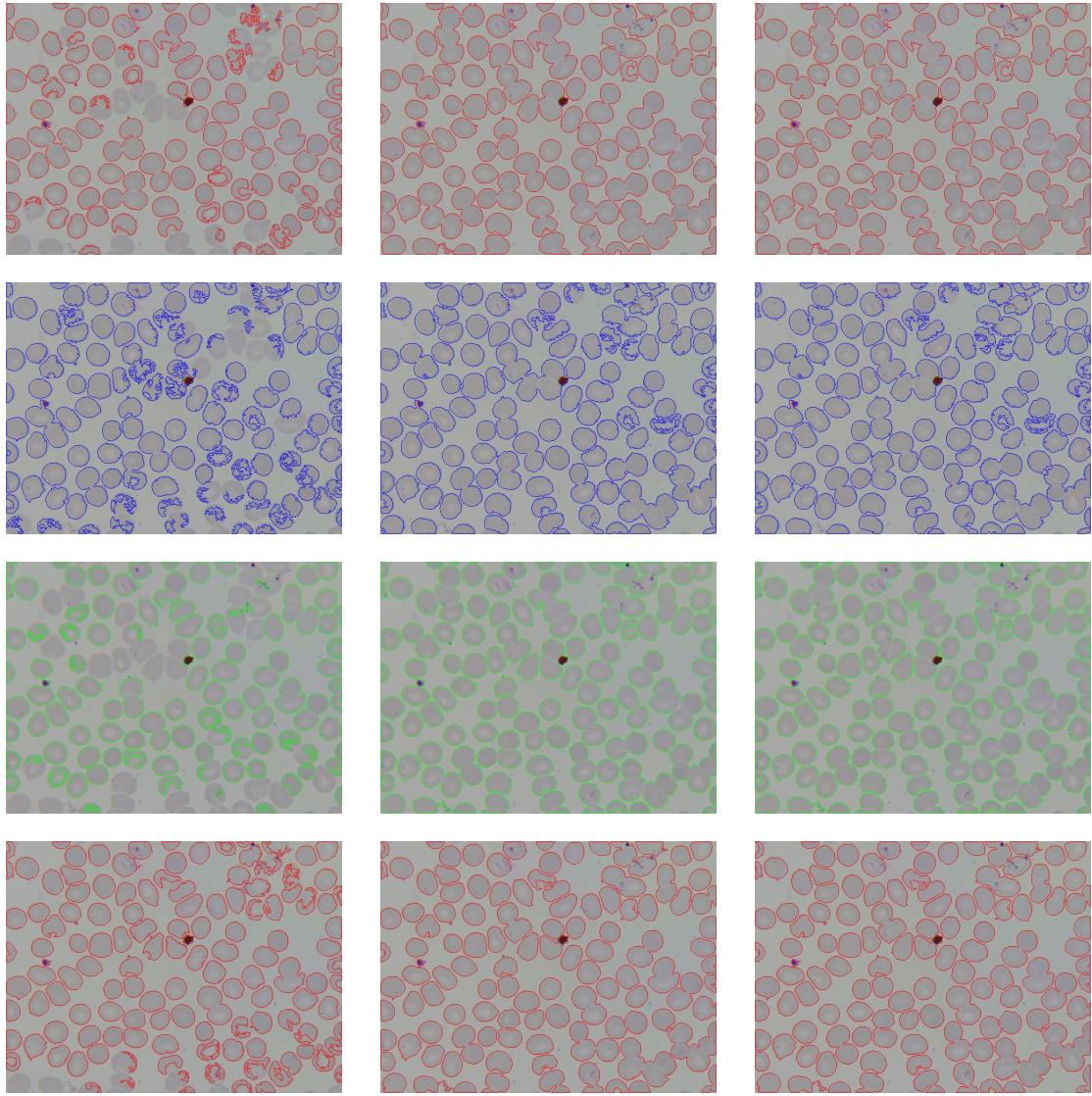


Area

Aspect ratio

Fill ratio

Figure 15: Segmentation results for separated cells (continued)

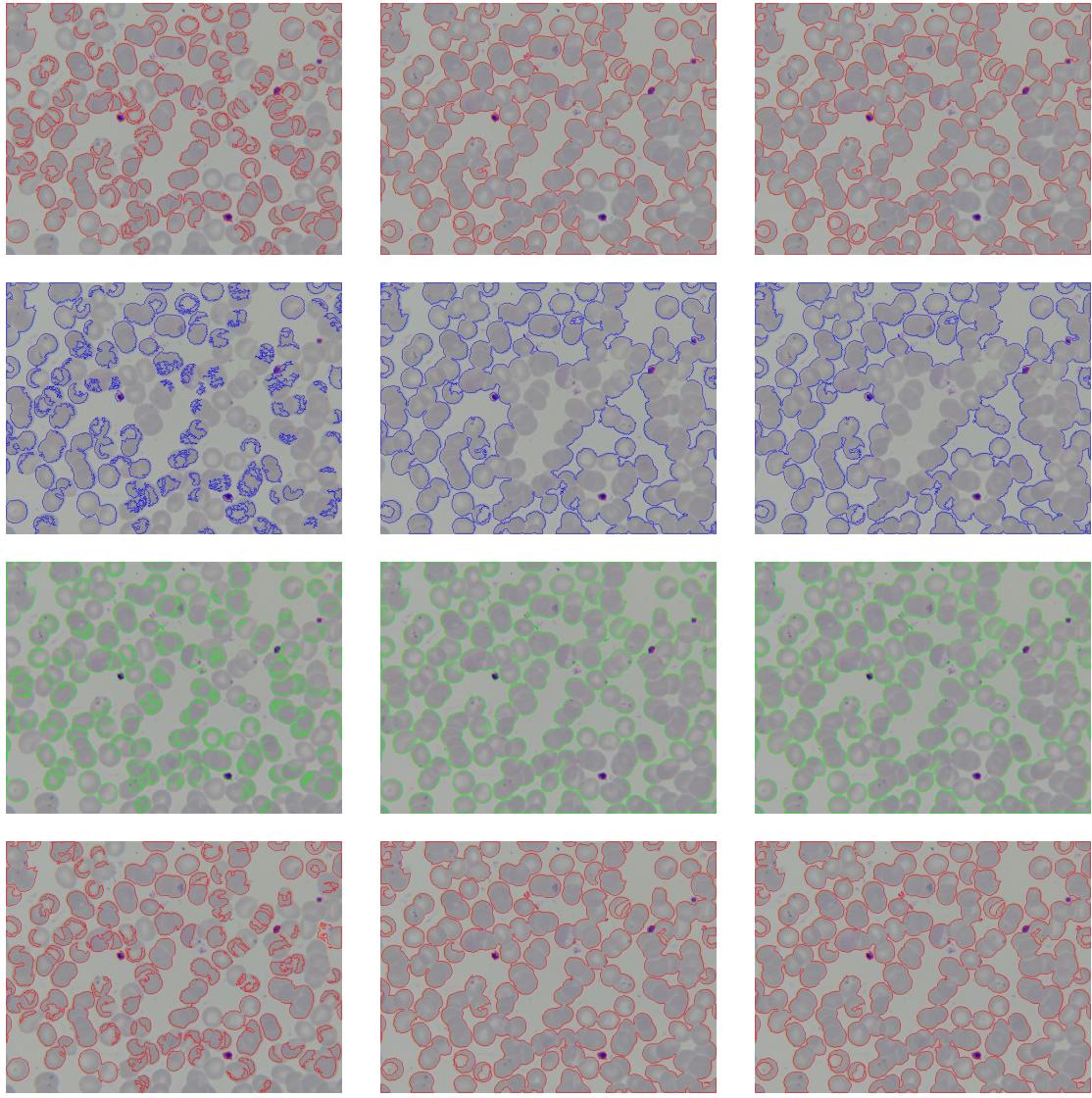


Area

Aspect ratio

Fill ratio

Figure 16: Segmentation results for separated cells (continued)

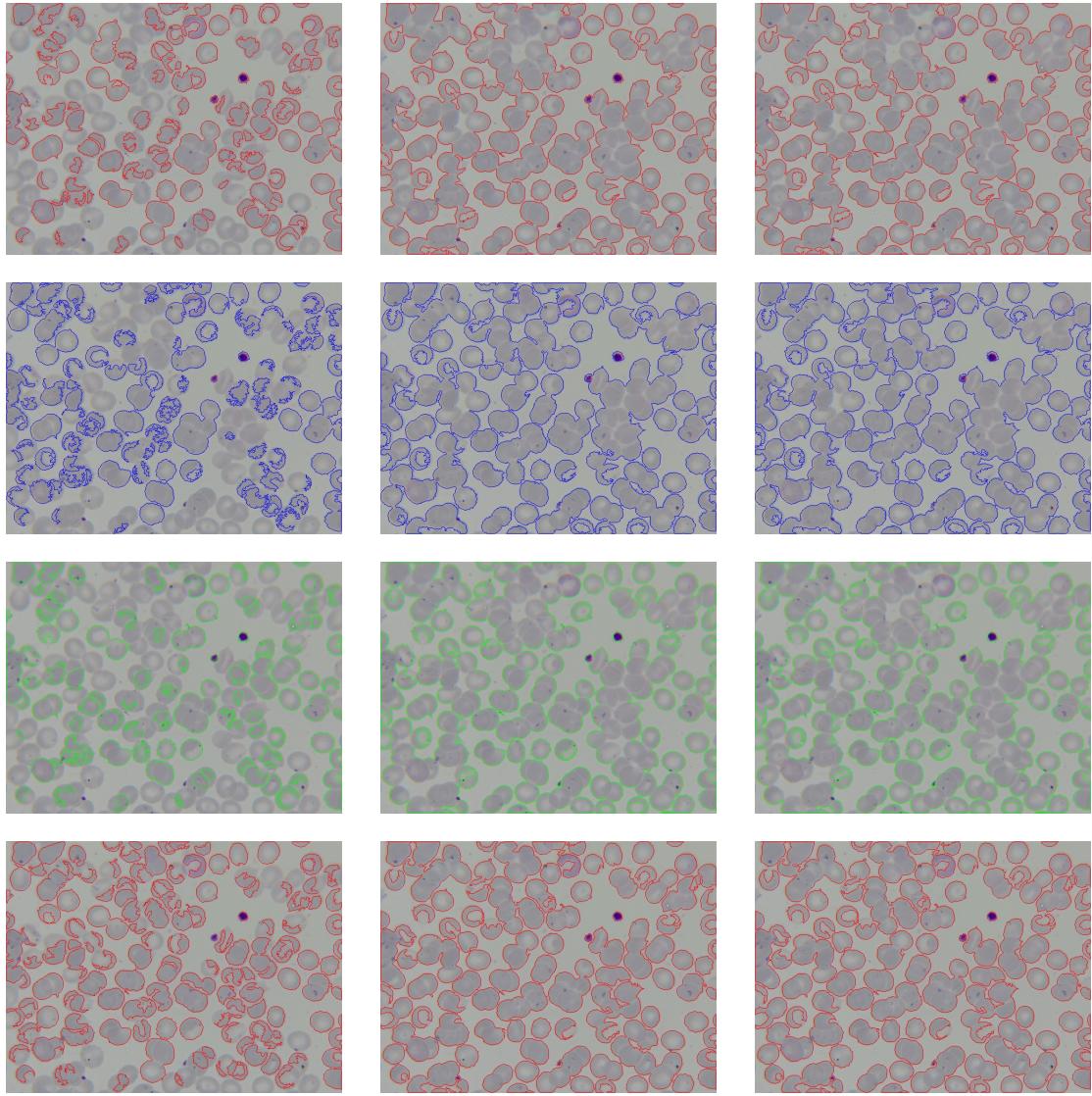


Area

Aspect ratio

Fill ratio

Figure 17: Segmentation results for overlapping cells

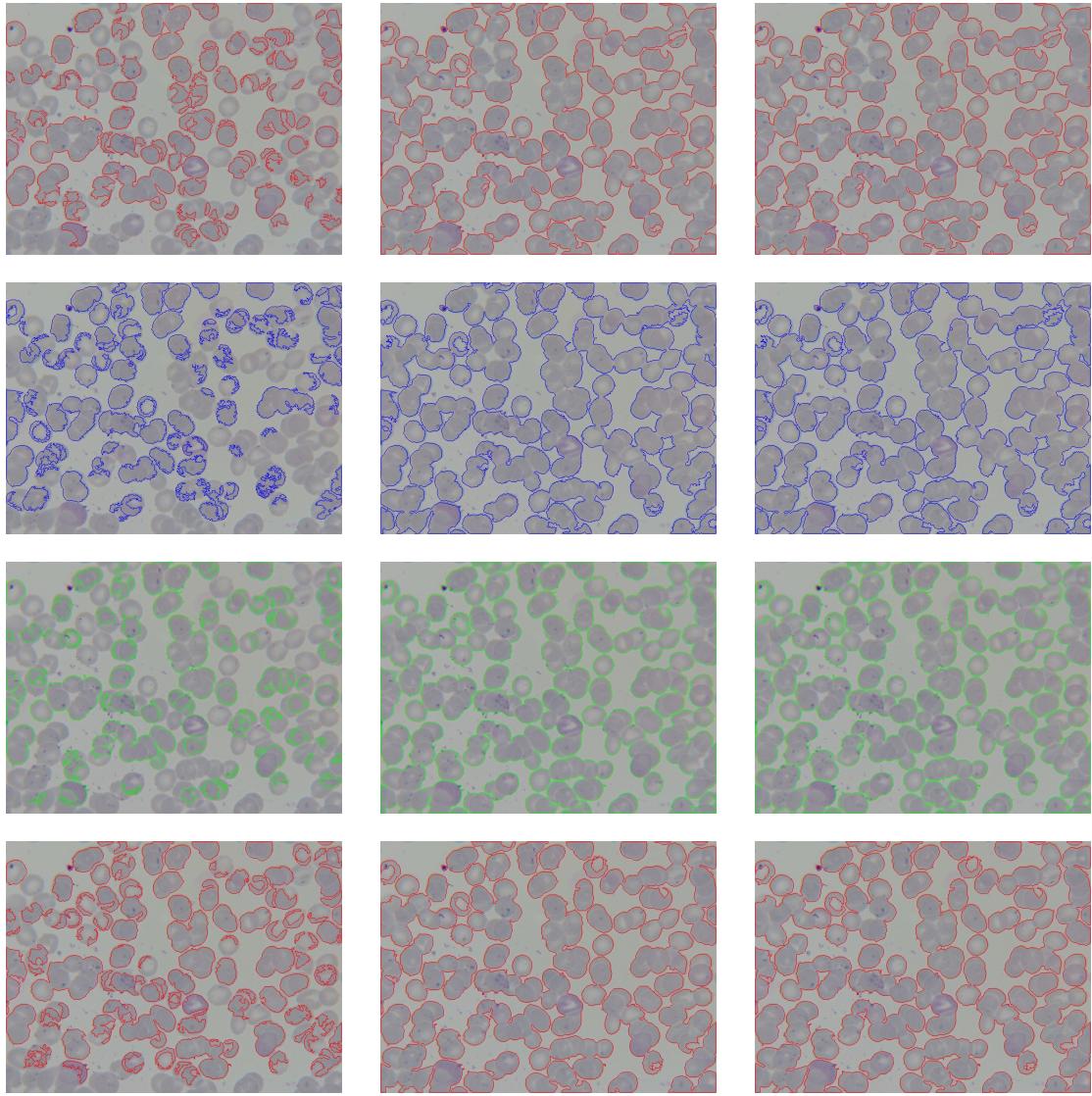


Area

Aspect ratio

Fill ratio

Figure 18: Segmentation results for overlapping cells (continued)

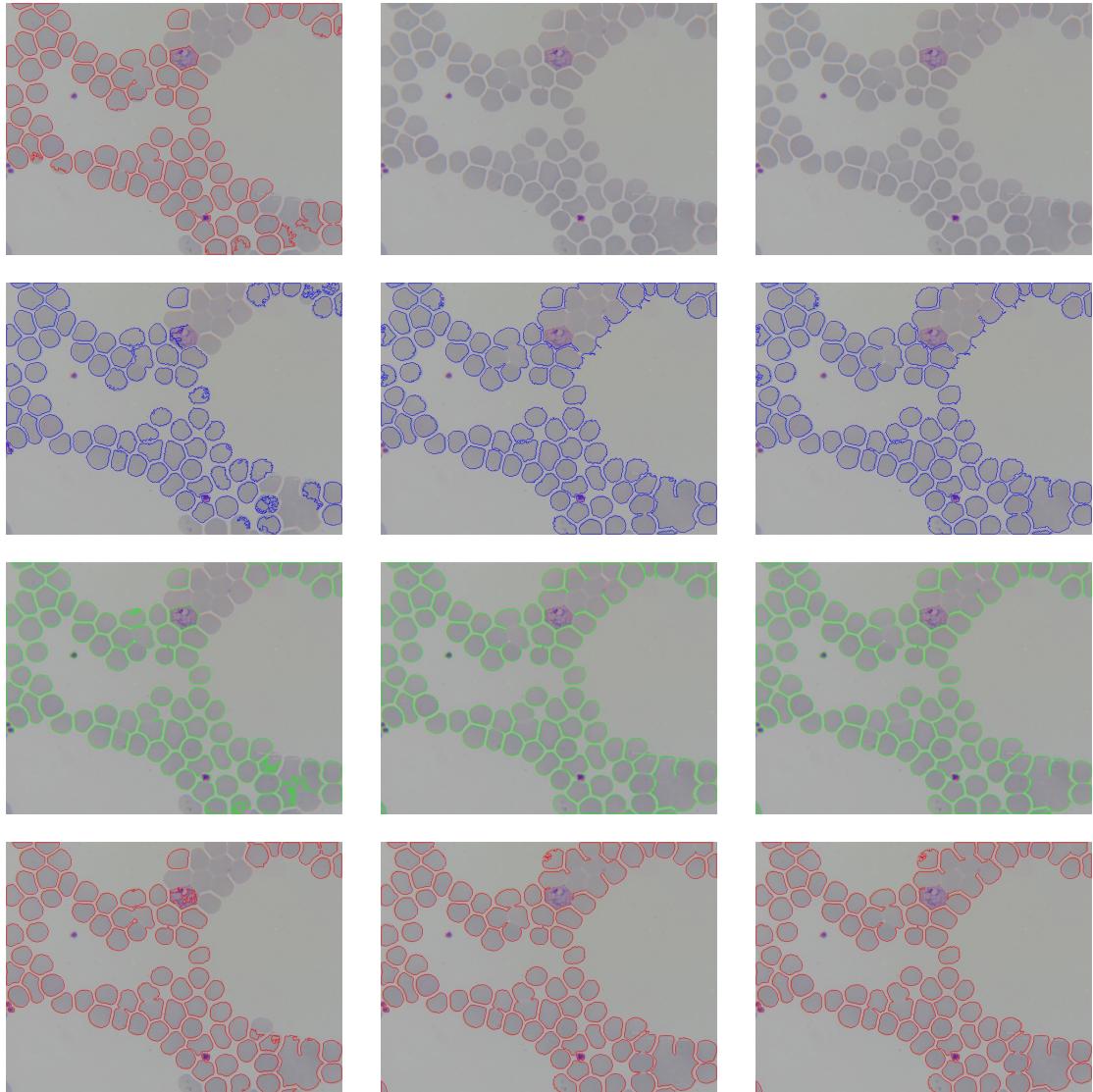


Area

Aspect ratio

Fill ratio

Figure 19: Segmentation results for overlapping cells (continued)

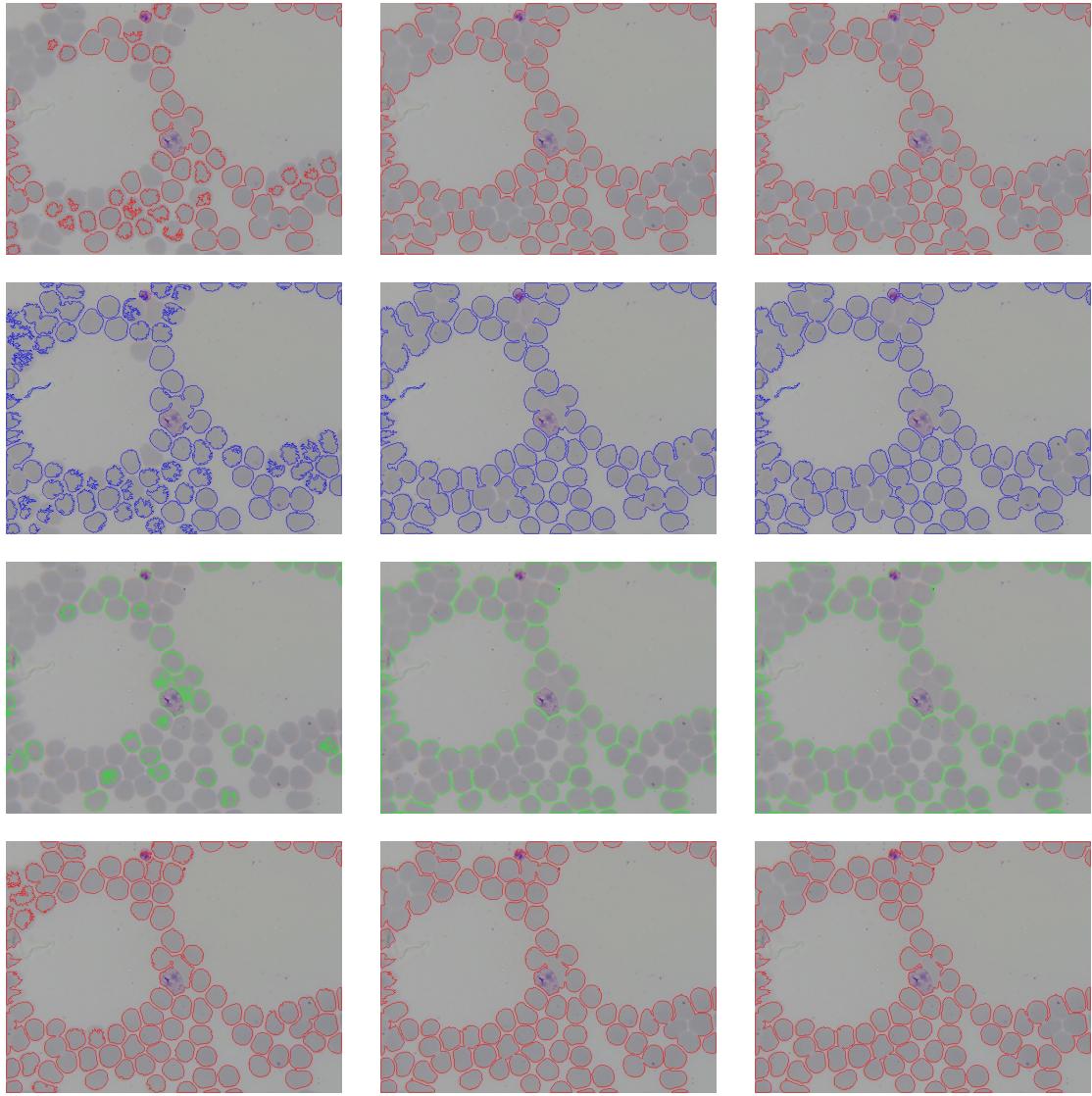


Area

Aspect ratio

Fill ratio

Figure 20: Segmentation results for squashed cells

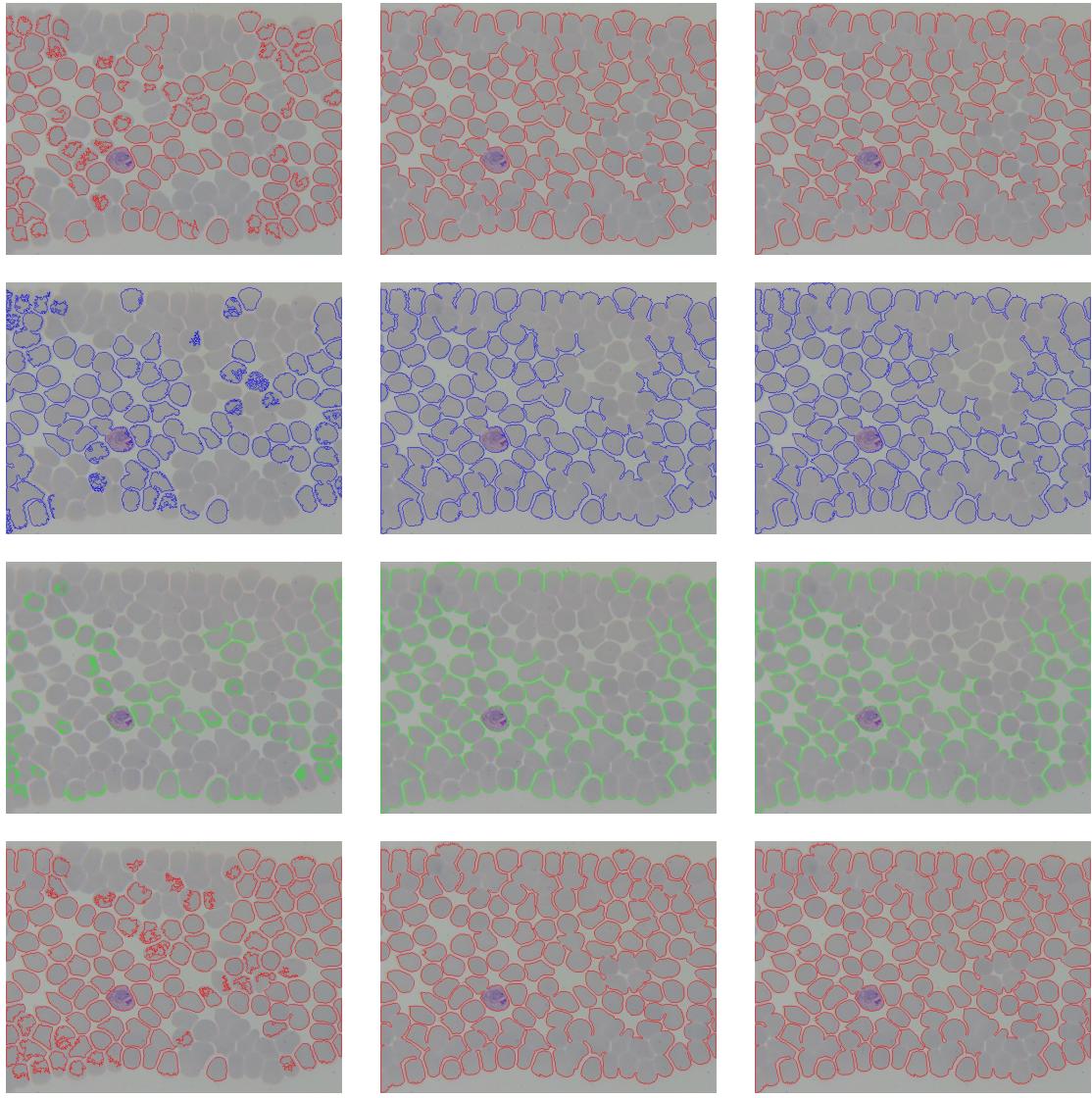


Area

Aspect ratio

Fill ratio

Figure 21: Segmentation results for squashed cells (continued)



Area

Aspect ratio

Fill ratio

Figure 22: Segmentation results for squashed cells (continued)

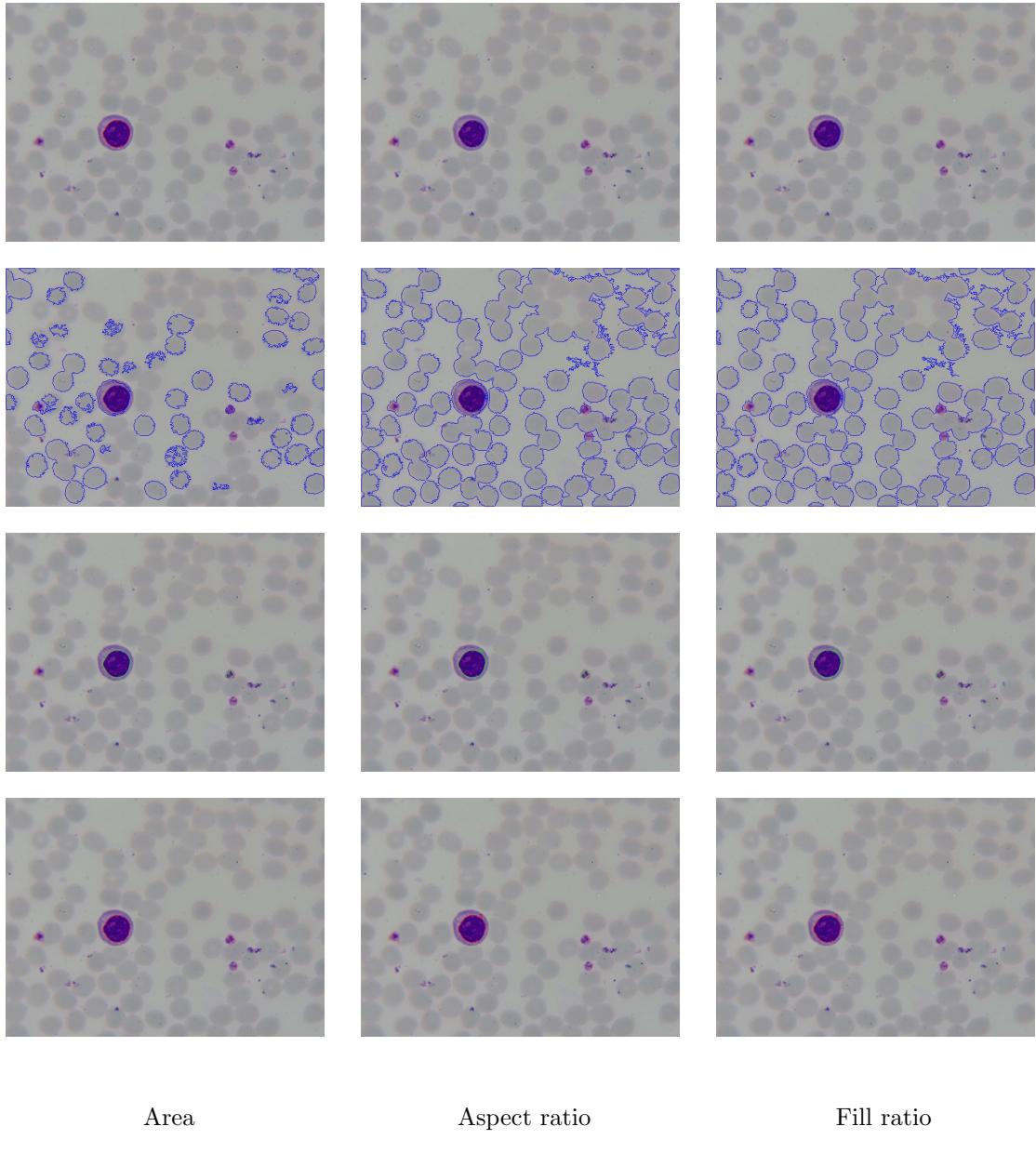
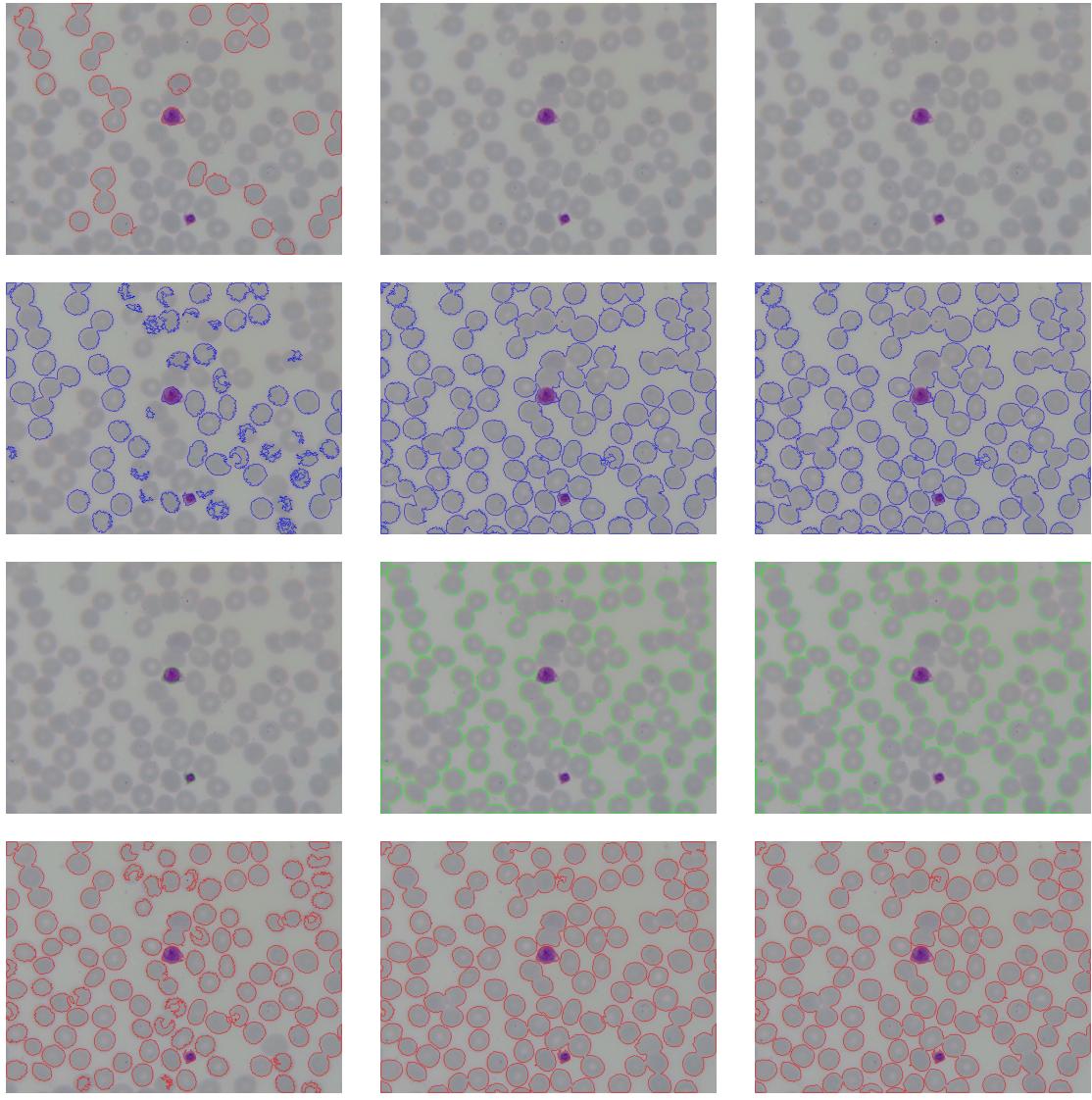


Figure 23: Segmentation results for blurred cells

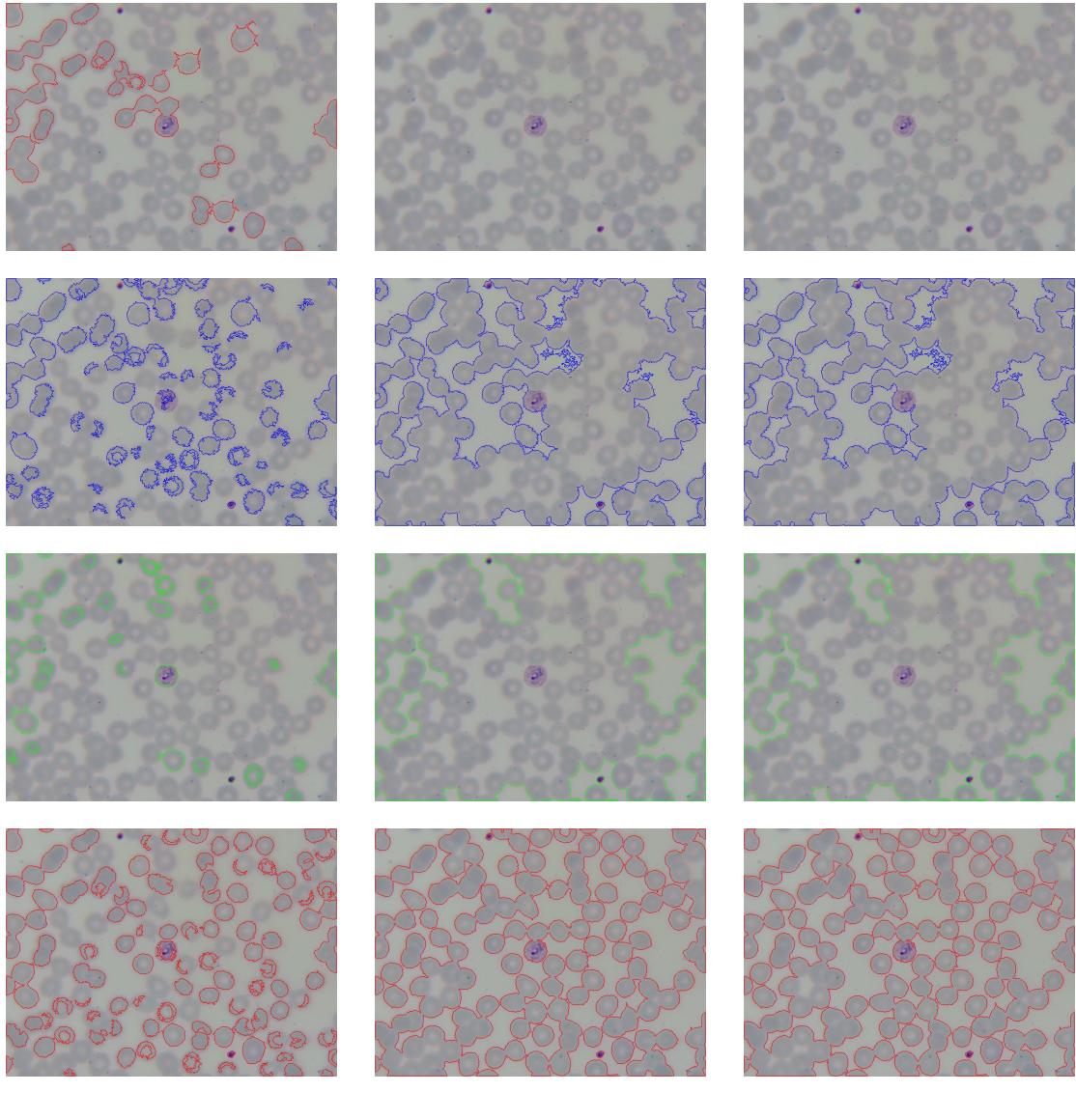


Area

Aspect ratio

Fill ratio

Figure 24: Segmentation results for blurred cells (continued)



Area

Aspect ratio

Fill ratio

Figure 25: Segmentation results for blurred cells (continued)