Segmentation of White Blood Cells Image Using Adaptive Location and Iteration

Yuehua Liu, Feilong Cao, Jianwei Zhao, and Jianjun Chu

***Abstract*—Segmentation of white blood cells (WBCs) image is meaningful but challenging due to the complex internal characteristics of the cells and external factors, such as illumination and different microscopic views. This paper addresses two problems of the segmentation: WBC location and subimage segmentation. To locate WBCs, a method that uses multiple windows obtained by scoring multiscale cues to extract a rectangular region is proposed. In this manner, the location window not only covers the whole WBC completely, but also achieves adaptive adjustment. In the subimage segmentation, the subimages preprocessed from the location window with a replace procedure are taken as initialization, and the GrabCut algorithm based on dilation is iteratively run to obtain more precise results. The proposed algorithm is extensively evaluated using a CellaVision dataset as well as a more challenging Jiashan dataset. Compared with the existing methods, the proposed algorithm is not only concise, but also can produce high-quality segmentations. The results demonstrate that the proposed algorithm consistently outperforms other location and segmentation methods, yielding higher recall and better precision rates.**

***Index Terms*—Adaptive adjustment, GrabCut algorithm, image segmentation, location, white blood cell (WBC).**

I. INTRODUCTION

A. Background

WITH the rapid development of computer image processing, the automatic detection, classification, and counting of white blood cells (WBCs) could be accomplished using microscopic images of WBCs. This automation is of great significance to increase efficiency and reduce a doctor's burden, as traditional manual counting is dull, tedious, and potentially subjective.

For given microscopic images, the counting work generally includes WBC segmentation, feature extraction, and classification. Among these, a key question is how to use "good" procedures to segment WBCs, because "good" procedures of segmentation often bring high accuracy of segmentation, and the accuracy will directly affect the medical diagnosis and therapy results.

However, WBC segmentation is challenging and complex work because of the various shapes and sizes of cells, uneven dying, cell adhesion, and some external environment factors, such as illumination and different microscopic views. Because of this, so far there are no fast and efficient methods to generate accurate results.

Though much remains to be done, results from recent studies show promise toward quicker more efficient WBC segmentation. Some general segmentation algorithms used for natural images, such as edge detection [1], automated thresholding [2], [3], morphology [4], region growing [5], multiscale analysis [6], and watershed [7], [8], have been modified and applied for the segmentation of WBCs. Considering the fact that the H and S components in colour space contain most of the information about the cell and nucleus, respectively, the method based on the HSI colour space is applied most commonly [9]. Wu *et al.* [10] segmented WBCs using iterative Otsu's approach based on the HSI colour space. The fuzzy clustering method and neural network technique were also used to divide the cell image into several regions in the segmentation in [11], [12], and [13], respectively.

Early in 1987, Kass *et al.* [14] proposed a type of image segmentation method called the "snake model," which initiated studies of the "active contour model" and has progressively replaced the conventional algorithm in segmentation nowadays [15].

This unhackneyed strategy has been used to extract WBCs from the output of curve evolution [15]. In recent years, a spectral segmentation approach using spectral graph theory [16] has become a new trend for natural images. This approach first uses the local information of a given image, and then partitions the image into meaningful image regions according to the global information computed from the affinity matrix. Of course, this method can also be applied to WBC segmentation [17]. However, the only drawback is that conducting spectral analysis of a large matrix requires an excessively expensive computation. All the aforementioned approaches suffer from interference from red blood cells (RBCs), plasma, platelets, and other surrounding factors, and also are sensitive to noise caused by illumination, dyeing, and so on, which will result in some lighter WBCs being ignored and darker RBCs being extracted. Rather than directly partitioning the input image into regions that we are interested in, object detection segmentation

---

Manuscript received June 7, 2016; revised October 4, 2016; accepted October 22, 2016. Date of publication November 4, 2016; date of current version November 3, 2017. This work was supported by the National Natural Science Foundation of China under Grant 61672477 and Grant 61571410. *(Corresponding author: F. Cao.)*

Y. Liu, F. Cao, and J. Zhao are with the College of Sciences, China Jiliang University, Hangzhou 310018, China (e-mail: 1281264596@qq.com; feilongcao@gmail.com; zhaojw@amss.ac.cn).

J. Chu is with the Jiashan Jasdaq Medical Device Corporation, Ltd., Jiashan 314100, China (e-mail: 1500641418@qq.com).

Digital Object Identifier 10.1109/JBHI.2016.2623421

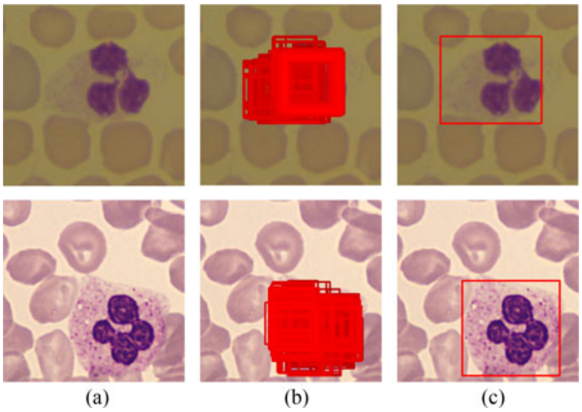


Fig. 1. Location window for WBC. (a) Original image. (b) Obtained multiple location window. (c) Final output location window.

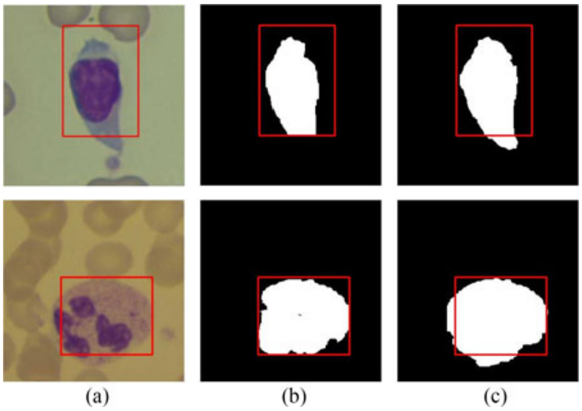


Fig. 2. Iterative GrabCut segmentation applied in a subimage (a) is the original image containing a red location window. Then, instead of manually drawing a box to initialize the GrabCut segmentation process, we automatically take the red window as initialization. The segmentation results from one iteration are shown in (b). In what follows, dilation and GrabCut segmentation proceed iteratively. Thus, some regions belonging to WBCs, while not detected previously, would eventually be segmented, as displayed in (c).

approaches [18], [19] segment, the object only from location subimages. This approach was motivated by the observation that subimages would assist in avoiding irrespective disturbs from surrounding factors and can achieve more promising results.

This object detection process was called "image size reduction" in [18], where the centre of each WBC nucleus was used to fit an appropriate window centered around it called the "subimage." Finally, the typical snake algorithm was made in the subimage for the subsequent segmentation.

B. Motivation and Contribution

In this paper, we present a new approach for WBC location and subimage segmentation. In general, object detectors are usually formulated as a single window or salient region to locate the object [19]–[22]. In our methods, however, a measure for locating WBCs is defined by multiple windows, and the detector follows the sliding window paradigms [23]–[26]. A single window would comprise only a very salient part of WBCs, i.e., regions not containing the lightly dyed cytoplasm, which will have a big influence on later feature extraction and further classification. In this way, our measure corresponds better to the whole and complete WBC. The following Fig. 1 illustrates our results for WBC location with defined measure using multiple windows. In what follows, the iterative GrabCut algorithm [19] is applied to the subimage containing the located WBC, and then the regions not detected precisely would be eventually segmented afterward. Fig. 2 shows the segmentation procedure. Compared with all the previous methods, our framework requires neither predetermined sizes of the location window nor network training, as in [20], but achieves better location performance.

We found that almost all WBCs have two characteristics [20]: 1) The density of edges in the edgemap produced using the Canny detector [27] is always concentrated in regions with WBCs, because of cells' natural characteristics. 2) The darker colours of WBCs differ from those of RBCs and the background. These characteristics can be used to define the detection measure.

Because WBCs have a variety of sizes, shapes, and locations, the traditional single sliding window used for detection does not apply to WBCs. Our multiwindow method, however, can cover the whole WBC.

The contributions of this paper can be summarized as follows:

1) Two distinctive characteristics of WBCs are utilized to select windows with higher scores. A new technique, using multiple windows, is then introduced to locate the WBCs. Unlike the previous single window with a fixed size [20], multiple windows allow adaptive adjustment of the location window, which will contribute to a better location result.
2) The idea of multiple scales is also explored. Before the multiple windows are integrated, fixed size sliding windows are scored. Rather than generating the multiple windows at a certain scale, we extend it to multiple scales. In particular, we make use of two fixed sizes to generate multiple windows in our experiment.
3) Ways to remove noisy windows are designed respectively for different measure cues. In addition, the output windows are integrated in a specific way considering the characteristics of the WBCs.
4) The subimage covering the WBC location is preprocessed for better use of the GrabCut algorithm, then an iterative GrabCut algorithm based on the dilation is applied to the segmentation of the whole WBC.
5) For the CellaVision [28] and Jiashan datasets, we demonstrate that the multiwindow measure performs better than other location methods, such as the location centered around the nucleus [18], [29], and the single fixed window location [20], and also, the next segmentation does better as well compared with the existing methods.

C. Organization

The remainder of this paper is organized as follows. Section II describes the WBC location in detail, including the measure cues, window generation, and multiwindow integration. The specific iterative segmentation based on the dilation is discussed in Section III. Specifically, we consider how to preprocess the subimage for better segmentation. Section IV shows

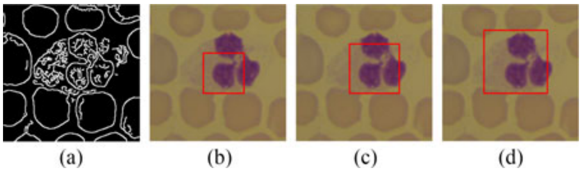


Fig. 3. Different sized windows with highest ED scores. (a) ED, (b)–(d) 90×90, 110×110, and 140×140 pixels. The edgemap demonstrates that the density of edges is concentrated on the WBC. Canny algorithms are implemented by the MATLAB function "edge". and the parameters are set by default values, with high threshold 0.7 and low threshold 0.28. The rest of the location images show location windows of different sizes with the highest ED scores.

the results and evaluations of location and segmentation, respectively. Finally, some conclusions and further remarks are made in Section V.

II. WBC LOCATION

In this section, we describe the cues for measuring the difference between WBCs and the background, introduce multiscale and multiwindow versions of these cues, and show how to integrate the windows.

A. Measure Cues

As described in Section I-B, we consider two characteristics of WBCs, and describe the corresponding measure cues: Edge Density (ED) and Colour Contrast (CC).

Edge Density. Alexe *et al.* [20] proposed an ED measure based on the density of edges in the edgemap, which is obtained using a Canny detector. A higher ED score favors a window with a large density. As introduced in Section I, the density of edges is always concentrated on the WBC. Thus, a higher ED score accordingly corresponds to a window containing a WBC (see Fig. 3). The ED measures the density near the window.

The nearby inner of the window w is denoted by $\mathrm{Near}(w)$ and the area of $\mathrm{Near}(w)$ is defined as

$$|\mathrm{Near}(w)| = \theta_{\mathrm{ED}}^2 |w| \tag{1}$$

where $\mathrm{Near}(w)$ is a rectangular ring obtained by shrinking the window w proportionally by a factor θ_{ED} in width and height, which is set experimentally. To be more concrete, $\mathrm{Near}(w)$ consists of the pixels, which are inside the window w and the most adjacent to the window w. $|\cdot|$ indicates the number of pixels. In what follows, the ED cue score is computed by

$$s_{\mathrm{ED}}(w) = \frac{\sum_{p \in \mathrm{Near}(w)} B(p)}{L(\mathrm{Near}(w))} \tag{2}$$

where the value in the binary edgemap B at each pixel p is 0 or 1, and $L(\mathrm{Near}(w))$ represents the perimeter of $\mathrm{Near}(w)$. Based on this cue, the following Fig. 3 shows that red windows score highest over all the slide windows in different sizes.

Colour Contrast. It is evident that the darker colour within the WBC region differs from that of the background and its surrounding area. Thus, the CC between the window and the surrounding area outside the window would also favor the window containing the WBC. The surrounding area of the window w is denoted by $|\mathrm{Out}(w)|$. Assuming the width and height is H, $|\mathrm{Out}(w)|$ can be expressed as $|\mathrm{Out}(w)| = \theta_{\mathrm{CC}}^2 H^2 - H^2 = (\theta_{\mathrm{CC}}^2 - 1)|w|$, where $\mathrm{Out}(w)$ is a rectangular ring obtained by enlarging the window w proportionally by factor θ_{CC} in width and height. This paper refers to [20] in defining $\mathrm{Out}(w)$ and choosing parameter. The CC measure is computed by the χ-square distance between the two histograms g and h of $\mathrm{Out}(w)$ and w based on Lab color space

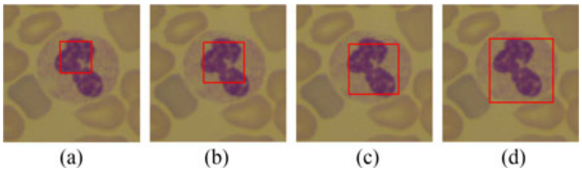


Fig. 4. Different sized windows with highest CC scores. (a)–(d) 90×90, 110×110, 140×140, and 170×170 pixels. For the CC detection, we also use a red colour for the highest score. The window with the highest score always covers the darkest colour which belongs to the WBC nucleus.

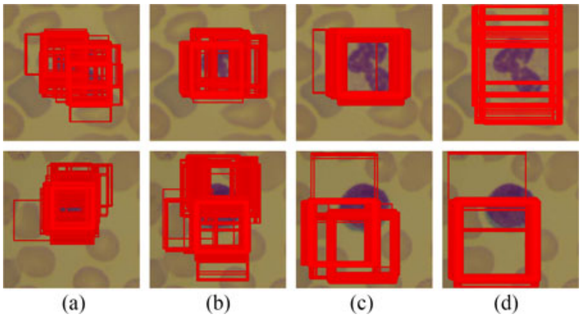


Fig. 5. Multiwindow ED based on multiple scales is better. (a)–(d) 90×90, 110×110, 140×140, and 170×170 pixels. Given different scales, windows scoring lower are rejected. The red windows represent those with higher scores. Whether the window scores in the first 3% is used as the distinguishing criterion. In terms of WBCs of varying sizes, we choose two small relative sizes to balance integrality and exactness.

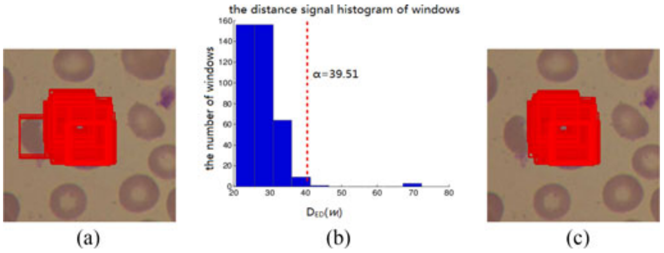
$$s_{\mathrm{CC}}(w) = \chi^2(g,h) = \frac{1}{2}\sum_i \frac{(g(i)-h(i))^2}{g(i)+h(i)} \tag{3}$$

where i represents the histogram index. In Fig. 4, the red window covering the WBC nucleus receives the highest score.

B. Multiwindow With Multiscale Cue

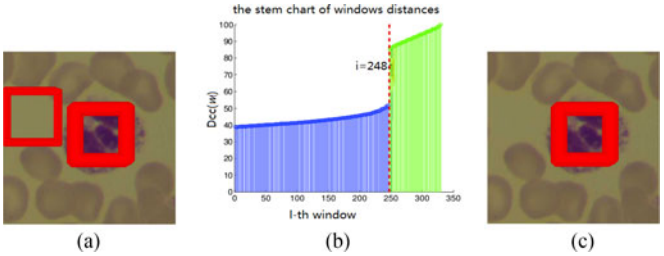
Because the size and location of a WBC within the image always appears different, the location window that can cover well the WBC should also vary in size. However, one slide window with a fixed size is unable to meet this requirement, we propose a new multiwindow method to achieve adaptation of the location window.

Multiscale ED Windows. Relying on the sliding window paradigm, our method selects the windows with the ED score ranked in the first 3% (see Fig. 5), and this threshold is set by experiment experience (thresholds in the range of about 2%–4% all could be used). To locate the whole WBC as completely and precisely as possible, we additionally stretch the ED sliding window using multiple scales (see Fig. 5). As shown in Fig. 4, bigger sizes like 110×110 and 140×140 pixels, can not locate the WBC very exactly. On the other hand, for window scales too small like 10×10 and 30×30 pixels, some



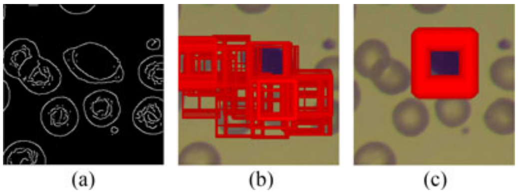
(a) (b) (c)

Fig. 6. Illustration of WBC localization using 50×50 pixels and results after the subsequent processing. (a) The location map with multiple windows in the first 3% ranked by ED score, (b) the histogram bar of the distance signal of those windows, and (c) location map after leaving windows to the right of the red dashed line in the middle histogram. Particularly, the red line, that is the threshold α, is computed via the rule. The windows with D_{ED} more than α would be removed.



(a) (b) (c)

Fig. 7. Illustration of localization based on CC cue. (a) The original location map, (b) the stem chart of those location windows, and (c) the final result after the removal operation. In the middle image, the first ith window satisfying the rule is 248th, so windows corresponding to the green point to the right of the red dashed line are removed.



(a) (b) (c)

Fig. 8. Stronger CC cue. (a) ED, (b) ED windows, and (c) CC windows. In a very special case, many Canny edges are not concentrated in the WBC region. (a) is such an example. A poor location result via the multiwindow method is shown in (b). However, the CC cue keeps its location box covering the WBC nucleus all the way.

RBC and blood platelet regions with high ED but small sizes, would be also extracted. So two square window scales are experimentally selected and set as 70×70 and 50×50 pixels in our experiment. (In fact, the window sizes between 40×40 and 90×90 all could be treated as reasonable selection). Moreover, when the microscope amplification changes, the size of WBCs images would appear different. The new scales should be reselected, and it is feasible to choose the scales, increased or reduced by the same proportion from 70×70 and 50×50, according to the ratio between different WBC sizes in different images. In our implementation, the windows are concentrated on the WBC, while there are still some noisy windows deviating from the WBC. Thus, some rules should be established to remove this noise. First, it is noted that the spatial distance between two windows is equipped with the Euclidean distance between their centroids. Moreover, for each window, we precompute the sum of the spatial distance with every other window, dividing by the total number as their distance metric. With M as the number of windows, the distance signal $D(w_l)$ of the l-th window w_l is computed by

$$D(w_l) = \sum_{j=1}^{M} \|x_l - x_j\|_2 \tag{4}$$

where x_l and x_j denote the coordinates of the window w_l and w_j centroids, respectively. A higher value of $D(w_l)$ indicates a greater possibility that the window w_l deviates from the WBC, where a majority of windows congregate. Thus, the windows with higher values would be labeled as outliers and need to be removed (see Fig. 6).

The rule proceeds by sorting and quantizing the $D(w_l)$ and removing the outliers using an adaptive threshold. Specifically:

1) Compute the distance signal $D(w_l)$ of each ED window, denote as $D_{\mathrm{ED}}(w_l)(l = 1, 2, 3, \ldots,)$;
2) Construct the histogram $(h(D_{\mathrm{ED}}))$ of $D_{\mathrm{ED}}(w_l)$. The distance signals of each window are sorted into some spaced bins along the X-axis from the minimum to maximum values $X_{\mathrm{ED}} = \{X_{\mathrm{ED}}(1), \ldots, X_{\mathrm{ED}}(r)\}$. Compute the height of each rectangle, which is the number of windows in the bin

$$N_{\mathrm{ED}} = \{N_{\mathrm{ED}}(1), \ldots, N_{\mathrm{ED}}(r)\}; \tag{5}$$

3) Remove less intensive windows. Find the first ith bin satisfying the following inequality, and remove windows with D_{ED} more than a threshold $\alpha = X_{\mathrm{ED}}(i + 1)$:

$$N_{\mathrm{ED}}(i + 1) < 0.2 N_{\mathrm{ED}}(i); \tag{6}$$

4) If the bin is not found, we tend to simply set α as the average of $D_{\mathrm{ED}}(w_l)$.

CC Windows. Analogously, the sliding window is chosen as 90×90 pixel (any size about from 80 to 110 is also available). Using the sliding window paradigm, windows whose CC score is in the top 3% are drawn (see Fig. 7). In special cases, the plasma with light colour may be located when it is surrounded by darker RBCs. In this way, their CC scores are also possibly high; therefore, removing this kind of window is desirable. Compared with windows detected by ED cue, windows detected by CC are more discrete (see Fig. 7). Thus, we directly use the simple method of distance sorting instead of histogram removal in the ED rule. Subsequently, a new rule aiming specifically at CC is designed as follows:

1) Compute the distance signal $D(w_l)$ of each CC window, and denote as $D_{\mathrm{CC}}(w_l)(l = 1, 2, 3, \ldots,)$.
2) Remove the outliers. Sort the $D_{\mathrm{CC}}(w_l)$, and find the first ith window satisfying $D_{\mathrm{CC}}(w_{i+1}) - D_{\mathrm{CC}}(w_i) > \beta$, where β is a distance threshold, simply set as 10 in our experiments.
3) Delete windows whose D_{CC} is higher than $D_{\mathrm{CC}}(w_i)$.

Multiwindow Integration. In the actual experiment, special circumstances arise in very few images. Fig. 8 shows that high CC is obviously stronger evidence for WBC location than ED. Thus, we integrate the windows mainly by CC as a criterion. Denoting the windows obtained by CC and ED as set C_i and E_i, respectively, the detailed algorithm process is represented in Algorithm 1.

Algorithm 1: WBC Adaptive Detection.

```
Input: WBC images ($f$);
Output: The location window of WBC;
1: Obtain multiwindows $E_i$ and $C_i$;
2: Set $I = \bigcup_i C_i$;
3: Calculate the overlap rate (OLR) (also known as
   Intersection over Union [30]) between $I$ and each
   window $E_i (i = 1, 2, 3, \ldots, )$;
4: Keep all windows $C_i$, and discard the windows $E_k$:
      $\{E_k \in E_i, (i = 1, 2, 3, \ldots, ) : \mathrm{OLR}(E_k) < \mu\}$,
   where $\mu$ is a constant threshold (see Fig. 9);
5: Find the window $w^*$ covering all windows $w^{(j)}$ from
   Step 4, $w^{(j)} \in C_i \cup E_i \setminus E_k, j = 1, 2, 3, \ldots$, Set
      $w^{(j)} = (tlx^{(j)}, tly^{(j)}, brx^{(j)}, bry^{(j)})$,
   where $(tlx^{(j)}, tly^{(j)})$ and $(brx^{(j)}, bry^{(j)})$ are
   coordinates of the top-left and bottom-right corner in
   the window $w^{(j)}$. We have
   $w^* = \Big(\min(tlx^{(j)}), \min(tly^{(j)}), \max(brx^{(j)}),$
          $\max(bry^{(j)})\Big)$
6: return The location window $w^*$.
```

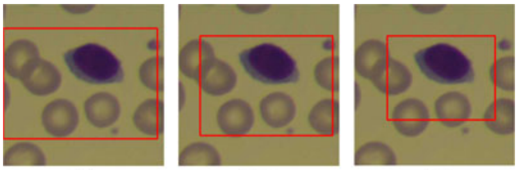

(a) (b) (c)

Fig. 9. Final location window by different threshold μ. (a)–(c) no threshold, $\mu = 0$, and $\mu = 0.1$.

III. WBC SEGMENTATION

The GrabCut algorithm [31] differs from the GraphCut [32] approach mainly in three aspects related to user experience. The first is that when users drag a rectangle roughly covering the region of interest, the object can be extracted well and automatically. Second, additional user interaction can improve the segmentation results. Third, a robust algorithm for "border matting" creates more precise and natural object segmentation boundaries. In this section, we introduce the iterative GrabCut algorithm. Instead of manually choosing a rectangle, the location window can be used as the initialization of the GrabCut algorithm to achieve automatic segmentation. In particular, we make a "dilate iterative" enhancement for some subimages not containing all WBC pixels.

A. Nucleus Segmentation

Nucleus always have the most distinctive color to the background, thus the nucleus should be extracted more easily. Actually there are already many approaches which can segment nucleus well, including [10],[13],[33]. Yet, when the algorithm is implemented based on the subimage; namely, the location window, they can do better, since the ill-effects from RBCs, platelets, and other factors have been eliminated. In our framework, the good segmentation of nucleus will have a great effect for the following preprocessing. But the key point of our paper is to explore the automatic and effective cytoplasm segmentation method. And here we would not given much discussion more further about nucleus segmentation.

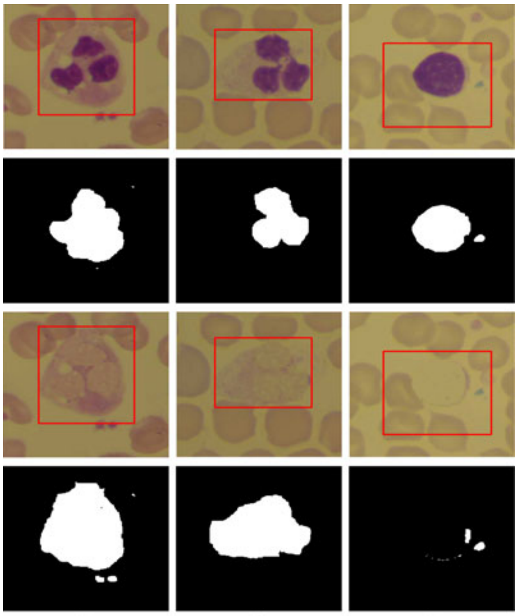

(a) (b) (c)

Fig. 10. Illustrations of three different examples. (a)–(b) successful examples, and (c) failure example. This is because of high contrast and subsequent success in finding cytoplasm pixels to substitute for the nucleus in the first two images. For the last image, however, there is too little cytoplasm, and plasma surrounding the nucleus was selected during the replace procedure.

B. Subimage Preprocessing

For cytoplasm segmentation, although we can simply cut out the WBC inside the subimage using the GrabCut algorithm, the entire WBC may not be completely segmented because of high contrast between the dark nucleus and light cytoplasm. Even in some *"bad"* (higher contrast of colour feature) images, the result unexpectedly tends to the nucleus (see Fig. 10(b)). To avoid these problems, we use a replace procedure to refine the subimage. The pixel values of the nucleus region are replaced by pixel values of the cytoplasm. We regard the regions surrounding the nucleus (NS) as representative pixels of the cytoplasm. Specifically, NS in our experiment is defined as

$$\mathrm{NS} = \mathrm{NU}_{\mathrm{dil}} - \mathrm{NU} \tag{7}$$

where NU, $\mathrm{NU}_{\mathrm{dil}}$, and NS represent the nucleus, dilated nucleus, and regions surrounding the nucleus, respectively. Then, for each pixel in the nucleus, the colours are randomly replaced by the colours in NS. A typical example of such a replace procedure is shown in the last two rows of Fig. 10. It should be noted that in cases where the WBC has little or no cytoplasm, the replace procedure may have the opposite effect (see Fig. 10(c)).

Algorithm 2: Iterative WBC Segmentation.

```
Input: $I_s$;
Output: Segmented WBC;
Initialize the GrabCut process by $I_s$, and set $i$ initially to 1;
repeat
    Dilate the current segmented WBC $I_s^{(i)}$, and obtain a new dilated region;
    Use the new region for the next GrabCut, and the running result is $I_s^{(i+1)}$;
    Set $i \leftarrow i + 1$;
until the $I_s^{(i)}$ converges to an accurate segmentation result.
return WBC $I_s^{(i)}$.
```

Thus, we set the size of the segment region after one iteration of the GrabCut algorithm as criterion to decide whether to preprocess the subimage. The larger region is considered the initial image for the subsequent iteration, while the smaller region is discarded. For each location subimage obtained from Section II, the selective principal is described as

$$I_s = \begin{cases} I_{np\_s}, & \text{if } |I_{np\_s}^{(1)}| \geq |I_{p\_s}^{(1)}| \\ I_{p\_s}, & \text{if } |I_{np\_s}^{(1)}| < |I_{p\_s}^{(1)}| \end{cases} \tag{8}$$

where I_s denotes the initial image used to iteratively run GrabCut for the final result, I_{p_s} and I_{np_s} represent the processed subimage and unprocessed subimage, and $I_{p\_s}^{(i)}$ and $I_{np\_s}^{(i)}$ are the segmented images after the ith iteration using I_{p_s} and I_{np_s} for the initialization, respectively.

C. Iterative GrabCut Segmentation

Once I_s is selected, we can iteratively use the GrabCut algorithm to improve the segmentation performance. We also add dilation operations on the current segmented WBC after each iteration. A new dilated region is then used for the next iteration. Thus, this iteration procedure enables the segmentation to include the nonlocated pixels (see Fig. 2). The cytoplasm iterative GrabCut segmentation scheme based on dilation is summarized in Algorithm 2, and overall work from location to segmentation is shown in Fig. 11. A comprehensive quantitative evaluation of different location methods and segmentation results is presented in Section IV.

IV. EXPERIMENTAL EVALUATION

A. Datasets and Setup

In our work, we evaluate our location and segmentation methods on the images from the CellaVision [28] and Jiashan datasets, both of which contain colour images. Some samples from two datasets would be given in Fig. 12. Our experimental tests include 60 color images collected from CellaVision reference library, which is also used in [29], containing normal images found in peripheral blood with a clear boundary, even dyeing and centred on the nucleus, and images are mostly at 360 × 360 resolution with five classes: Neutrophil, Lymphocyte, Monocyte, Eosinophil, and Basophil.

For more extensive evaluations, the Jiashan dataset is also used. The dataset is collected from The First People's Hospital of Jiashan, Zhejiang Province, China, and contains more than 800 images with 500 × 500 resolution. Note that these images are gathered directly without preselection, so eosinophils and basophils account for only a small part of the dataset, much like the cell distribution in human blood. However, much of our focus should be on the challenges posed by images in this dataset, because some images may have ambiguous boundaries, lighter cytoplasm, irregular shape, and off-centre nucleus.

Our work targets WBC location and segmentation, its performance and efficiency should be compared with those of other algorithms. The WBC images from CellaVision have been already segmented and classified by experts within the field of hematology, the ground truth is included in dataset. For Jiashan dataset, we invited the experts from The First People's Hospital of Jiashan to make the human segmentation as the ground truth. Specifically, in the following experiments, the ground truth for location is just the circumscribed rectangle of human segmented WBCs added with acceptable error $0 - 9$ pixels. In summary, the human annotations serve as the ground truth in the form of a rectangle for location and binary map for segmentation. All the location and segmentation algorithms are running on a PC with CPU and a 1.90 GHZ Inter i7-3517U processor using our hybrid programming with MATLAB code, OpenCV, and C language.

B. Benchmarks

This paper mainly concentrates on the WBC location and subimage segmentation. While there are many approaches for these purposes, a method that measures the quality of the location and segmentation is desirable.

F-Measure [34]. The F-measure is introduced for the purpose of evaluating a test's performance and is described by the precision-recall curve, which is a parametric curve that captures the tradeoff between accuracy and error. The precision-recall curve is a standard evaluation technique in the information retrieval community [34], and was first used in evaluating edge detectors [35]. It considers both the precision P and the recall R to compute the F score. Denoting expert and algorithm segmentation as G and I, respectively, we have

$$P = \frac{G \cap I}{I}, R = \frac{G \cap I}{G}. \tag{9}$$

Precision is assessed by computing the ratio of true positives to the all positives found by the proposed algorithm; namely, the machine segmentation, while recall is the percentage of true positives to all real positives; that is, the ground truth segmentation. In what follows, the F score can be interpreted as the harmonic average between precision and recall

$$\frac{\beta^2 + 1}{F} = \frac{1}{P} + \frac{\beta^2}{R} \text{ such that } F = \frac{(\beta^2 + 1)PR}{\beta^2 P + R} \tag{10}$$

where β is a harmonic parameter.

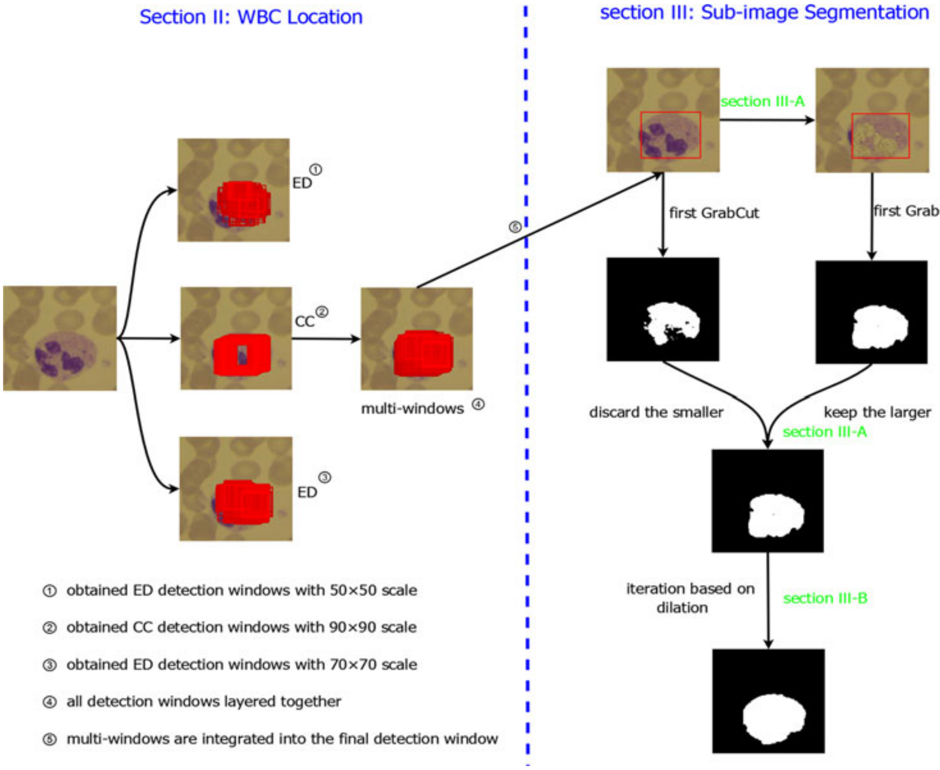


Fig. 11. Workflow of our adaptive location and iterative segmentation based on dilation.

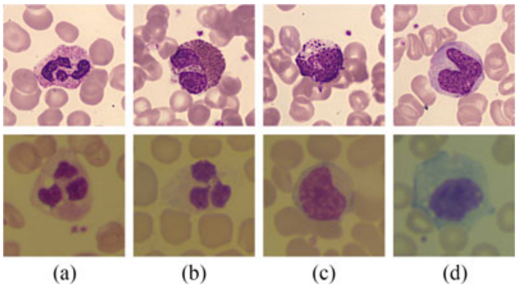


Fig. 12. Some samples from CellaVision [28] in first row and Jiashan datasets in last row.

Overlap [36], [37]. The overlap between the regions R and R' is defined as

$$O(R, R') = \frac{|R \cap R'|}{|R \cup R'|}. \tag{11}$$

In [36] and [37], an overlap metric is introduced to propose the segmentation covering, which is expressed as

$$C(S' \to S) = \frac{1}{N} \sum_{R \in S} |R| \max_{R' \in S'} O(R, R') \tag{12}$$

where $C(S' \to S)$ represents the covering of expert segmentation S by algorithm segmentation S', and N is the number of pixels in the image. However, as for our particular case, unlike the natural images with an ambiguous boundary, the WBC segmentation result is a certain and clear region (see Fig. 13). Thus, compared with segmentation covering, the overlap metric is sufficient as an accuracy measure.

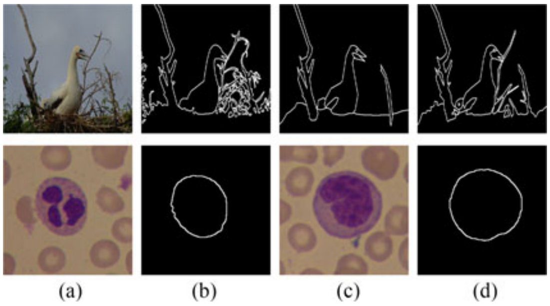


Fig. 13. Berkeley segmentation dataset (BSDS500) [38] and WBC ground truth segmentation. In the first row, (a) natural image, and (b)–(d) the ground truth segmentation boundaries by three different experts. In the bottom row, (a), (c) WBC images, and (b), (d) corresponding segmentation ground truths. We can observe that the WBC segment ground truth boundary has a definite criteria, independent of expert subjectivity.

C. Location Quality

We compare the proposed algorithm with two typical methods used in [18], [20], and [29]. Although the size of the location window in [18] is chosen as 141×141, it should vary in different datasets. To reliably compare the performance of different location methods, as for the method in [18], 140×140, 150×150, 160×160, and 170×170 are all used to fit the appropriate window, respectively. Visual comparisons of location subimages are presented in Fig. 14. The method used in [18] clearly has a poor effect on WBC images not centred on the nucleus, and the detection method in [20] is not appropriate for WBC images at all. The WBC that is too loose and not quite

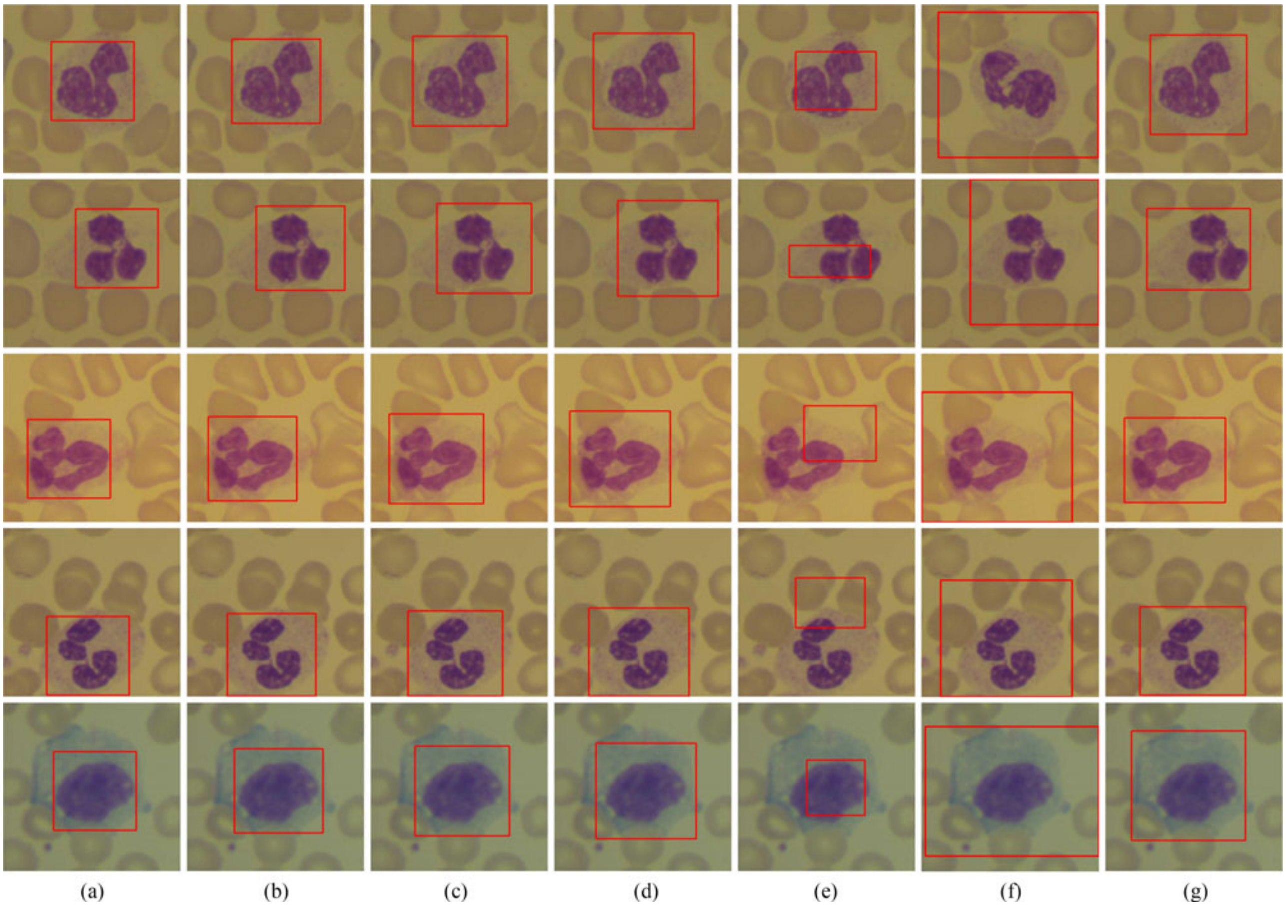


Fig. 14. Visual comparisons of different location subimages. (a) [18] with 140×140 pixels, (b) [18] with 150×150 pixels, (c) [18] with 160×160 pixels, (d) [18] with 170×170 pixels, (e) Alexe [20], (f) Ko [29], and (g) our method. Our method achieves adaptive adjustment of windows and generates the most precise location windows.

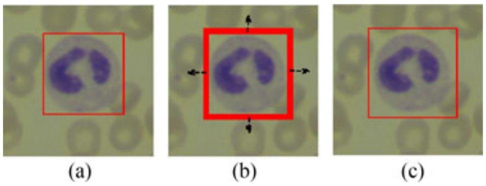


Fig. 15. Ground truth examples. (a) the ground truth with the added pixels 0, which is exactly the exterior rectangle of an expertly segmented WBC, (b) the ground truth with the added pixels from 0 to 9 in the directions of the arrows, and (c) the ground truth with the added pixels 9.

precise is always framed in [29]. Our approach, however, has satisfactory results.

For a comprehensive quantitative evaluation, the ground truth criterion should be defined first. Subjectivity and some manual errors are inevitable in expert location. Thus, we use the exterior rectangle of an expertly segmented WBC with pixels added in four directions as the location ground truth (see Fig. 15). To comprehensively compare how well different methods locate WBCs, we then vary the added pixels from 0 to 9 to obtain the change in overlap rate curves (see Fig. 16).

The overlap rate curves clearly demonstrate that our multiwindow location method outperforms other methods. We observed that the overlap rate in [18] with 140×140 and 150×150 and our method trend downward. In contrast, those in [18] with 160×160 and 170×170 show an upward trend. This is because the location windows obtained from the former approaches are relatively smaller, the numerator $|R \cap R'|$ does not change much while the denominator $|R \cup R'|$ becomes larger along the X-axis (added pixels). However, the location map obtained by the latter methods is bigger, the numerator becomes larger while the denominator remains nearly the same. Overall, for the added pixels among 1–9, which is the exact range for human annotation error, our method is obviously better than the others.

The bottom in Fig. 16 shows that statistical comparison results by F-measure based on the CellaVision and Jiashan datasets. Unlike attention detection in [19] and [39], we look more at WBC integrity while considering the precision. In other words, we use $\beta^2 = 5$ to weigh recall more than precision, while the method in [19] uses $\beta^2 = 0.3$ to weigh precision more. For the CellaVision dataset, our approach reduced 33.7%, 25.8%, 21.7%, 20.1%, and 70.7% overall error rates on F-measure, compared with [18] with 140×140, 150×150, 160×160, 170×170, and [20], respectively. Similarly, for the Jiashan dataset, 20.5%, 16.4%, 16.3%, 18.9%, and 74.7% error rates are reduced. The two arrows point out that the algorithm in [18] would sacrifice the precision in order to achieve a high recall rate, with a larger location window. Thus, we can see from Fig. 16 that the methods in [18] with 140×140 and 150×150

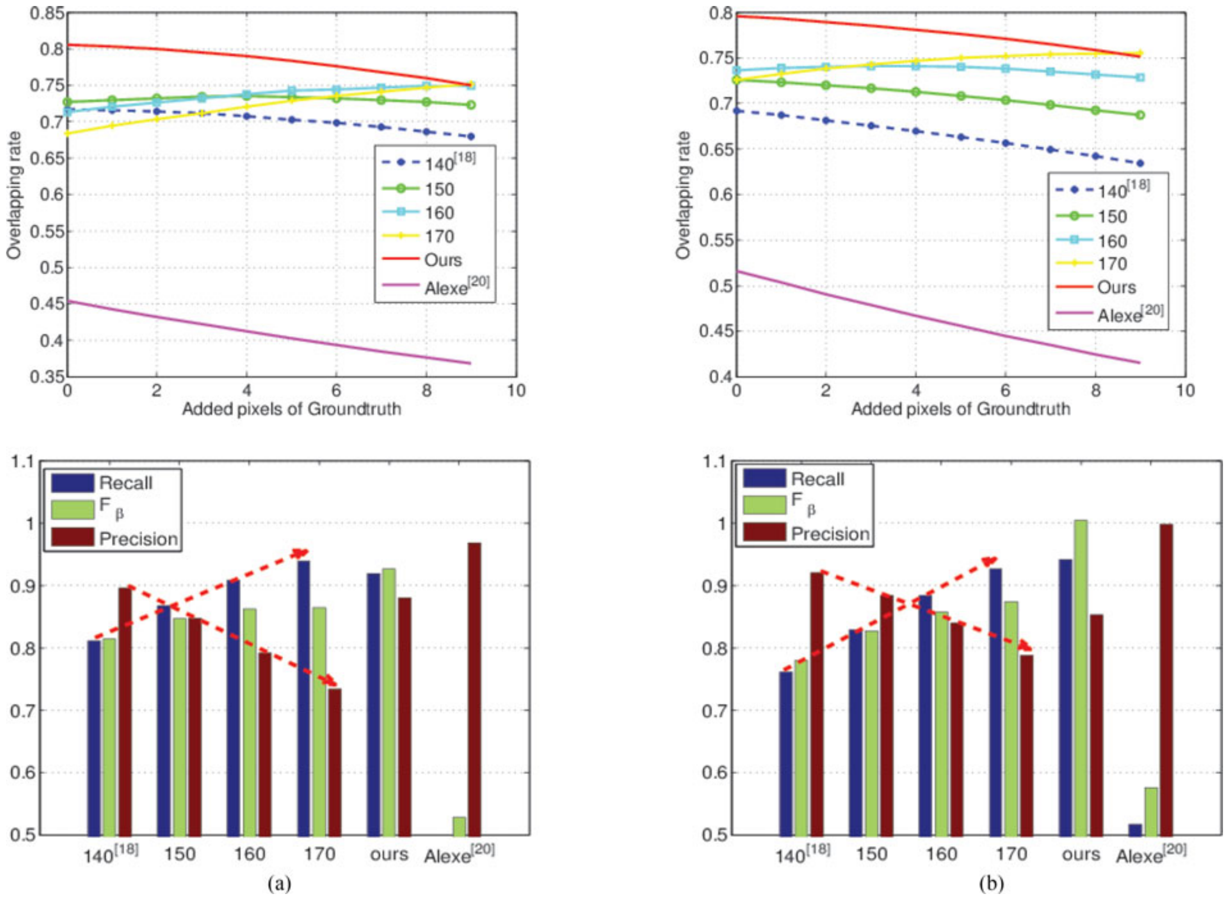


Fig. 16. Statistical comparisons of different location subimages. The first row is their overlap rate and subsequent evaluation by F-measure, using Jiashan and CellaVision datasets, respectively. (a) Jiashan dataset. (b) CellaVision dataset [28]

have a high precision but low recall rate, while approaches in [18] with 160 × 160 and 170 × 170 have a high recall rate but low precision, and our method achieves relatively high precision and recall rate. This means that in the actual experiment, the former methods can not contain the whole WBC, while the latter method tends to produce larger location rectangles. Our approach is, by contrast, much more precise.

The visual comparisons in Fig. 14 and evaluations in Fig. 16, all show the contributions and great significance of using multiple windows. At the same time, better locations will have a big influence on the subsequent segmentation procedure.

TABLE I
NUMERICAL COMPARISONS OF DIFFERENT SEGMENTATION RESULTS BASED ON JIASHAN DATASET

Methods	Mohapatra [13]	Ko [29]	Wu [10]	Tareef [40]	Ours
Precision	0.8021	0.7831	0.9784	0.8419	0.9331
Recall	0.7941	0.9108	0.6309	0.6955	0.8793
F_β	0.7688	0.7979	0.8407	0.7660	0.9164

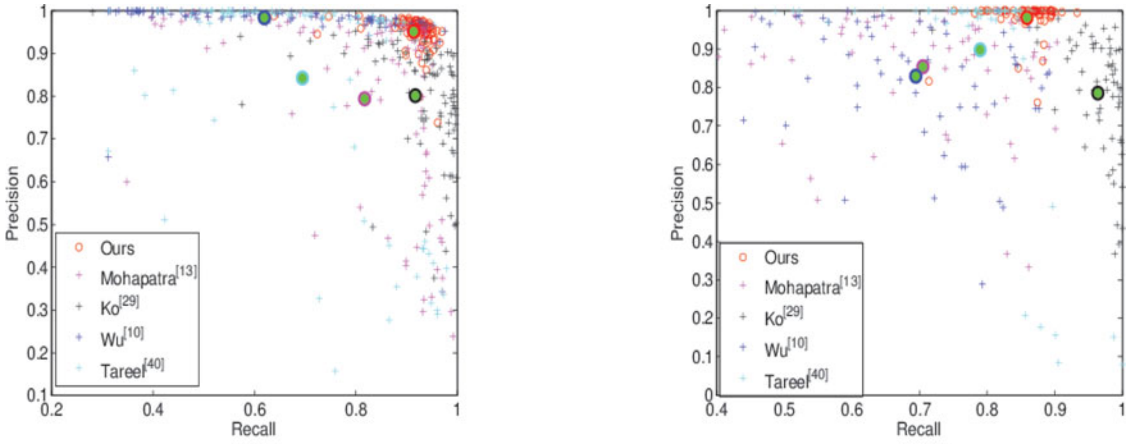
TABLE II
NUMERICAL COMPARISONS OF DIFFERENT SEGMENTATION RESULTS BASED ON CELLAVISION DATASET

Methods	Mohapatra [13]	Ko [29]	Wu [10]	Tareef [40]	Ours
Precision	0.8534	0.7849	0.8293	0.8981	0.9818
Recall	0.7049	0.9637	0.6944	0.7897	0.8592
F_β	0.8001	0.8138	0.7793	0.8459	0.9499

D. Segmentation Quality

In this section, we compare our method with some typical segmentation methods in [10], [29], and [13], and one up-to-date paper [40], then evaluate the results based on the CellaVision dataset and the challenging Jiashan dataset, which can comprehensively illustrate how well different methods segment WBCs. In theory, the key question regarding the gradient vector flow curves in [29] is the overflow of edges, leading to the sharp spike. This would have a big influence on the subsequent classification procedure from the circularity cue. Neither the threshold method in HSI spaces in [10] nor the neural networks classification in [13] would handle the complicated background well, either the darker RBCs would be segmented, or the lighter plasma of WBCs would not be segmented. Moreover, in [40], the decorrelation stretch, discrete wavelet transform, and morphological filtering is utilized in enhancement procedure. Then cytoplasm is extracted by Otsu's thresholding. But this algorithm also appears not to work in challenging images with very light dyeing.

All above analysis would be verified visually from the comparison shown in Fig. 18. Results from a quantitative evaluation are shown in Fig. 17 and Tables I and II. Fig. 17 illustrates the



(a) (b)

Fig. 17. Statistical comparisons of different segmentation results. Precision-recall values are displayed in the scatter plots based on (a) randomly selected images from Jiashan, and (b) CellaVision datasets, respectively. Solid dots with the different colour correspond to their average values. (a) Jiashan dataset. (b) CellaVision dataset [28].

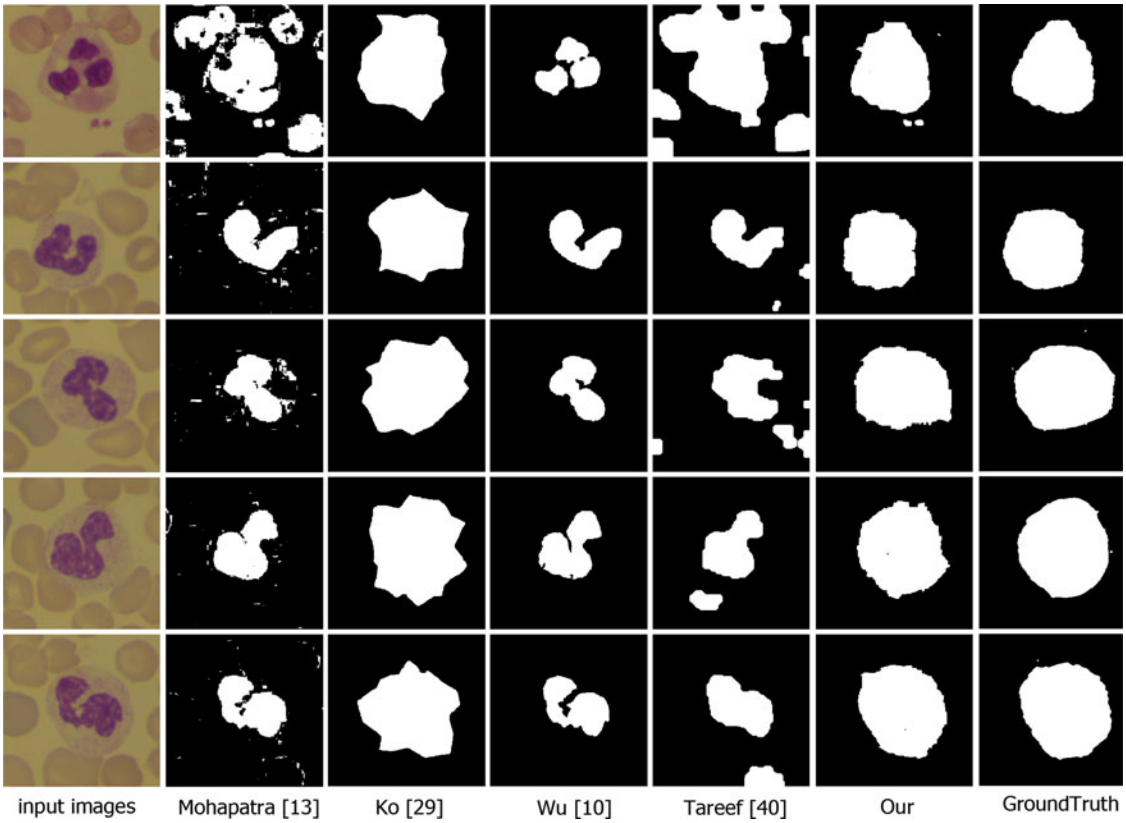


Fig. 18. Visual comparisons of different segmentation results for challenging WBCs.

segmentation performance per image, and from that we can see that our results present better and more stable. In Tables I, II, the detailed numerical results from different methods are given, Experiments indicate that some WBCs segmented via other methods may have either low recall or precision rate, but our method has both high recall and precision rate. Our average precision-recall value is the highest and much closer to (1,1). Thus, compared to the other methods, ours gives more promising results. This success is attributed to better location and iteration based on dilation. Moreover, we also make a comparison on each of five cell types in Table III. Since Jiashan dataset does not have the calibration of classified information, we only use the CellaVision to experiment. But images from CellaVision almost all have the clear boundary and even dyeing, so the segmentation performance is not surprisingly about same in different cell types. In general, basophil should have the highest precision and recall, and monocyte has the lowest precision and recall because of its wide area of cytoplasm and even same color with RBCs. However, the bad influences from RBCs have been eliminated after location procedure. Then,

TABLE III
SEGMENTATION PERFORMANCE COMPARISON ON EACH CELL TYPE

Cell Type	Neutrophils	Eosinophils	Basophils	Lymphocytes	Monocytes
Precision	0.9870	0.9982	0.9605	0.9680	0.9956
Recall	0.8531	0.8481	0.8393	0.8659	0.8895
F_β	0.9519	0.9587	0.9289	0.9410	0.9688

Results are obtained based on the CellaVision dataset.

from Fig. 12(d) in the first row, we can observe that the monocyte image from CellaVision has the fairly clear boundary and distinguishing color. So, based on the accurate location window, GrabCut algorithm can be initialized to segment monocyte well. In contrast, some basophils in CellaVision (see Fig. 12(c) in the top row), have lighter cytoplasm, which is actually not easy to be segmented.

V. CONCLUSION AND FURTHER REMARKS

In this paper, we present a new model of WBC segmentation, which first detects and then segments to achieve the accurate WBC segmentation. In detection stage, a multiwindow method based on two characteristics is proposed. We also select sliding windows with multiscales to fit the WBCs with different sizes. This method can be used to implement the adaptive adjustment of WBC location window. In the segmentation stage, once the location subimages are determined, an iterative GrabCut algorithm based on dilation is used to produce the final segmentation. Compared with the segmentation directly based on raw input images, the proposed framework can effectively avoid bad influences from background factors such as RBCs and platelets. In addition, different rectangular windows were fitted for different WBCs, whereas windows obtained by other methods were fixed squares with the predetermined size. Therefore, our location results are more accurate. Moreover, it only takes 0.6 s to locate WBC and 0.9 s to segment WBC in a testing image in CellaVision dataset on our experiment PC, and that can be ignored in human senses.

However, the proposed methods still have some disadvantages. For example, some parameters are experimentally set and their values within the appropriate range are all reasonable. Some questions, such as the existence of the optimal value of the parameters, and how to select and adjust them based on different datasets, are worth further study. In addition, the proposed algorithm only takes peripheral WBCs images into account. There is almost no adhesion between WBCs in peripheral blood, and in fact the ratio of WBCs to RBCs is about 1:1000 there. Actually no overlapping of WBCs happens in our experimental images. In my knowledge, severe adhesions are probably only found in marrow blood. Moreover, in general, the big-sized images, first taken under the microscope, contain hundreds of cells. Then, small-sized WBC regions are roughly cropped manually, containing one WBC. All images tested in our experiment just contain one WBC. So, we do not consider the cases of marrow WBCs, overlapped WBCs, and images with several WBCs included. This is one of the future works to be continued.

There are several remaining issues for future work on WBC classification. Since stronger identifiability features and a better classifier will further improve the classification performance, in the future, we expect to experiment with feature extraction from segmented WBCs, by manual selection or automatic extraction.

REFERENCES

[1] B. R. Kumar, D. K. Joseph, and T. V. Sreenivas, "Teager energy based blood cell segmentation," in *Proc. Int. Conf. Digit. Signal Process.*, 2002, vol. 2, pp. 619–622.

[2] I. Cseke, "A fast segmentation scheme for white blood cell images," in *Proc. 11th IAPR Int. Conf. Pattern Recognit., Conf. C: Image Speech Signal Anal.*, Aug. 1992, vol. III, pp. 530–533.

[3] Q. Liao and Y. Deng, "An accurate segmentation method for white blood cell images," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2002, pp. 245–248.

[4] C. Di Ruberto, A. Dempster, S. Khan, and B. Jarra, "Analysis of infected blood cell images using morphological operators," *Image Vis. Comput.*, vol. 20, no. 2, pp. 133–146, Feb. 2002.

[5] J.-M. Chassery and C. Garbay, "An iterative segmentation method based on a contextual color and shape criterion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 6, pp. 794–800, Nov. 1984.

[6] L. B. Dorini, R. Minetto, and B. J. Leite, "Semiautomatic white blood cell segmentation based on multiscale analysis," *IEEE J. Biomed. Health Infor.*, vol. 17, no. 1, pp. 250–256, Jan. 2013.

[7] L. B. Dorini, R. Minetto, and N. J. Leite, "White blood cell segmentation using morphological operators and cale-space analysis," *Brazilian Symp. Comput. Graph. Image Process.*, Oct. 2007, pp. 294–304.

[8] K. Jiang, Q.-M. Liao, and S.-Y. Dai, "A novel white blood cell segmentation scheme using scale-space filtering and watershed clustering," in *Proc. Int. Conf. Mach. Learn. Cybern.*, Nov. 2003, vol. 5, pp. 2820–2825.

[9] H. Hengen, S. L. Spoor, and M. C. Pandit, Analysis of blood and bone marrow smears using digital image processing techniques, *Proc. SPIE*, vol. 4684, pp. 624–635, 2002.

[10] J. H. Wu, P. Zeng, Y. Zhou, and C. Olivier, "A novel color image segmentation method and its application to white blood cell image analysis," in *Proc. 2006 8th Int. Conf. Signal Process.*, 2006, vol. 2.

[11] E. Montseny, P. Sobrevilla, and S. Romaní, "A fuzzy approach to white blood cells segmentation in color bone marrow images," in *Proc. IEEE Int. Conf. Fuzzy Syst.*, Jul. 2004, vol. 1, pp. 173–178.

[12] H. Sheikh, B. Zhu, and E. Micheli-Tzanakou, "Blood cell identification using neural networks," in *Proc. IEEE 22nd Annu. Northeast Bioeng. Conf.*, Mar. 1996, pp. 119–120.

[13] S. Mohapatra, D. Patra, S. Kumar, and S. Satpathy, "Lymphocyte image segmentation using functional link neural architecture for acute leukemia detection," *Biomed. Eng. Lett.*, vol. 2, no. 2, pp. 100–110, Jun. 2012.

[14] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, Jan. 1988.

[15] F. Zamani and R. Safabakhsh, "An unsupervised GVF snake approach for white blood cell segmentation based on nucleus," *In Proc. 8th Int. Conf. Signal Process.*, 2006, vol. 2.

[16] F. R. Chung, *Spectral Graph Theory*, vol. 92. Providence, RI, USA: American Mathematical Society, 1997.

[17] M. Oger, P. Belhomme, J. Klossa, J.-J. Michels, and A. Elmoataz, "Automated region of interest retrieval and classification using spectral analysis," *Diagnostic Pathol.*, vol. 3, no. 1, pp. 1–4, Jul. 2008.

[18] S. H. Rezatofighi and H. Soltanian-Zadeh, "Automatic recognition of five types of white blood cells in peripheral blood," *Comput. Med. Imag. Graph.*, vol. 35, no. 4, pp. 333–343, Jun. 2011.

[19] M.-M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu, "Global contrast based salient region detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 569–582, Feb. 2015.

[20] B. Alexe, T. Deselaers, and V. Ferrari, "Measuring the objectness of image windows," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2189–2202, Nov. 2012.

[21] S. Ravishankar, A. Jain, and A. Mittal, "Multistage contour based detection of deformable objects," *Eur. Conf. Comput. Vis.*, vol. 5302, pp. 483–496, Oct. 2008.

[22] B. Leibe and B. Schiele, “Scale-invariant object categorization using a scale-adaptive mean-shift search,” *Pattern Recognit.*, vol. 3175, pp. 145–153, Aug. 2004.

[23] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, vol. 1, pp. 886–893.

[24] P. Felzenszwalb, R. Girshick, D. McAllester, and D. Ramanan, “Object detection with discriminatively trained part-based models,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1627–1645, Sep. 2010.

[25] H. Harzallah, F. Jurie, and C. Schmid, C,, “Combining efficient object localization and image classification,” in *Proc. Int. IEEE 12th Conf. Comput. Vis.*, Sep. 2009, pp. 237–244.

[26] C. H. Lampert, M. Blaschko, and T. Hofmann, “Beyond sliding windows: Object localization by efficient subwindow search,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.

[27] J. Canny, “A computational approach to edge detection,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, no. 6, pp. 679–698, Nov. 1986.

[28] [Online]. Available: http://www.cellavision.com. Accessed: 2015.

[29] B. C. Ko, J.-W. Gim, and J.-Y. Nam, “Automatic white blood cell segmentation using stepwise merging rules and gradient vector flow snake,” *Micron*, vol. 42, no. 7, pp. 695–705, Oct. 2011.

[30] D. Yoo, S. Park, J.-Y. Lee, A. Paek, and I. S. Kweon, “Attentionnet: Aggregating weak directions for accurate object detection,” *IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 2659–2667.

[31] C. Rother, V. Kolmogorov, and A. Blake, “Grabcut: Interactive foreground extraction using iterated graph cuts,” *ACM Trans. Graph.*, vol. 23, no. 3, pp. 309–314, Aug. 2004.

[32] Y. Boykov, O. Veksler, and R. Zabih, “Fast approximate energy minimization via graph cuts,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, Nov. 2001.

[33] V. P. Ananthi and P. Balasubramaniam, “A new thresholding technique based on fuzzy set as an application to leukocyte nucleus segmentation,” *Comput. Methods Programs Biomed.*, vol. 134, pp. 165–177, Oct. 2016.

[34] C. Van Rijsbergen, *Information Retrieval*. Glasgow, Scotland: Univ. of Glasgow, 1979.

[35] I. E. Abdou and W. K. Pratt, “Quantitative design and evaluation of enhancement/thresholding edge detectors,” *Proc. IEEE*, vol. 67, no. 5, pp. 753–763, May 1979.

[36] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, “Contour detection and hierarchical image segmentation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, Mar. 2011.

[37] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, “From contours to regions: An empirical evaluation,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 2294–2301.

[38] D. Martin, C. Fowlkes, D. Tal, and J. Malik, “A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics,” in *Proc. 8th IEEE Int. Conf. Comput. Vis.*, Jul. 2001, vol. 2, pp. 416–423.

[39] T. Liu *et al.*, “Learning to detect a salient object,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 2, pp. 353–367, Feb. 2011.

[40] A. Tareef, Y. Song, W. Cai, Y. Wang, D. D. Feng, and M. Chen, “Automatic nuclei and cytoplasm segmentation of leukocytes with color and texture-based image enhancement,” in *Proc. 13th IEEE Int. Symp. Biomed. Imag.*, Apr. 2016, pp. 935–938.

Yuehua Liu received the B.Sc. degree in applied mathematics from China Jiliang University, Hangzhou, China, in 2014. She is currently working toward the M.Sc. degree in applied mathematics at China Jiliang University.

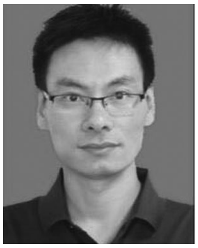
Her research interests include image processing, pattern recognition, and machine learning.

Feilong Cao received the B.Sc. and M.Sc. degrees in applied mathematics from Ningxia University, Yinchuan, China, in 1987 and 1998, respectively, and the Ph.D. degree in applied mathematics from Xi’an Jiaotong University, Xi’an, China.

He was a Research fellow with the Center of Basic Sciences, Xi’an Jiaotong University from 2003 to 2004. From 2004 to 2006, he was a Postdoctoral Research Fellow with the School of Aerospace, Xi’an Jiaotong University. From June 2011 to December 2011, and October 2013 to January 2014, he was a Visiting Professor in the Department of Computer Science, Chonbuk National University, Jeonju, South Korea, and the Department of Computer Sciences and Computer Engineering, La Trobe University, Melbourne, Vic, Australia, respectively. He is currently a Professor and the Dean of the College of Sciences, China Jiliang University, Hangzhou, China. He has authored or co-authored more than 150 scientific papers in refereed journals. His current research interests include pattern recognition, neural networks, and approximation theory.

Jianwei Zhao received the B.Sc. and M.Sc. degrees in mathematics from Shannxi Normal University, Xi’an, China, in 2010 and 2003, respectively, and the Ph.D. degree in mathematics from the Chinese Academy of Sciences, Beijing, China, in 2006.

She was a Postdoctoral Research Fellow in the Department of Electronics and Information Engineering, Chonbuk National University, Jeonju, South Korea, in 2011. She is currently a Professor and the Head of the Department of Applied Mathematics, College of Sciences, China Jiliang University, Hangzhou, China. Her current research interests include pattern recognition, machine learning, and neural networks.



Jianjun Chu was born in July, 1974. He received the Ph.D degree in optical engineering from the Beijing Institute of Technology, Beijing, China, in 2002.

He is currently the General Manager of Jiashan Jasdaq Medical Device Corporation, Ltd., Jiashan, China. His current research interests include biomedical optics, flow cytometry and medical image processing.