An Automated Blood Cell Detection and Segmentation System

Ligong Han\textsuperscript{a}, Ngan T.H. Le\textsuperscript{b}, and Marios Savvides\textsuperscript{c}

\textsuperscript{a}Department of Biomedical Engineering, Carnegie Mellon University
\textsuperscript{b, c}Department of Electrical and Computer Engineering, Carnegie Mellon University

ABSTRACT

This paper presents an end-to-end framework for automatically detecting and segmenting blood cells including normal red blood cells (RBCs), connected RBCs, abnormal RBCs (i.e. tear drop, burr cell, helmet, etc.) and white blood cells (WBCs). Our proposed system contains several components to solve different problems regarding RBCs and WBCs. We first design a novel blood cell color representation which is able to emphasize the RBCs and WBCs in separate channels. Template matching technique is then employed to individually detect RBCs and WBCs in our proposed representation. In order to automatically segment the RBCs and nuclei from WBCs, we develop an adaptive level set-based segmentation method which makes use of both local and global information. The detected and segmented RBCs, however, can be a single RBC, a connected RBC or an abnormal RBC. Therefore, we first separate and reconstruct RBCs from the connected RBCs by our suggested modified template matching. Shape matching by inner distance is later used to classify the abnormal RBCs from the normal RBCs. Our proposed method has been tested and evaluated on different images from ALL-IDB\textsuperscript{9}, WebPath\textsuperscript{24}, UPMC\textsuperscript{23} and the one used by Mohamed et al.\textsuperscript{14} The precision and recall of RBCs detection are 98.43\% and 94.99\% respectively, whereas those of WBCs detection are 99.12\% and 99.12\%. The F-measure of our proposed WBCs segmentation gets up to 95.8\%.

1. INTRODUCTION

Visual analysis of two-dimensional (2D) pathology images provides quantitative information about the presence and absence of disease process, and helps diagnosis of disease progress\textsuperscript{11,4}. In automated diagnosis systems, though human intervention can be necessary, it is desirable that the amount of intervention is minimum. Moreover, researches have shown that with improved segmentation accuracy, better diagnosis performance can also be achieved.\textsuperscript{5}

Some early methods for detection and segmentation were based on morphological operations. In paper by Kovalev\textsuperscript{8}, filtration of nucleus pixel template was utilized for localization, while region growing was used for segmentation and extraction of nucleus blobs. Similarly, morphological operation was also applied to remove red blood cells (RBCs) and segment white blood cells (WBCs) in an alternative approach.\textsuperscript{18} As for classification, manually designed morphological features were extracted and fed into classifiers including Naive Bayes classifier, k-nearest neighbors (k-NN), and multi-layer perceptrons (MLP). In Hiremath’s paper,\textsuperscript{3} a similar strategy was adopted, where the author performed erosion, reconstruction and dilation, followed by global thresholding. Despite their success, the algorithms mentioned above all operate solely on gray scale images. There have been researchers who exploited more information of color space. Ruberto’s paper in 2013\textsuperscript{20} adopted CMYK (cyan, magenta, yellow and key) color model for membrane detection, the result of which was fed to thresholding and Zack algorithm. The author also applied watershed segmentation to distance transform for separating adjacent leukocytes. Another approach by Mohapatra\textsuperscript{15} used k-means in the standard RGB (red, green, and blue) color space for cell detection, and shadowed C-means (SCM) in Lab color space for segmentation. A nucleus region enhancement in RGB color space was also proposed,\textsuperscript{19} where ellipse curve fitting was combined for segmentation. In a recent paper,\textsuperscript{2} the author utilized mean-shift and support vector machine (SVM) in RGB space for detection and segmentation respectively. Also, Liu\textsuperscript{11} proposed in his latest paper a novel algorithm which was performed in polar coordinate was proposed, and nuclei segmentation was then transformed into a Dijkstra shortest path problem. In our work, we focused mainly on blood cell detection and segmentation.

We observed two potential limitations in the methods mentioned above. First, many of these approaches rely on converting color images into gray scale images, which may induce a loss of color information. Even though some algorithms perform in RGB, Lab, CMYK, or other enhanced color spaces, the effectiveness may vary from case to case. Second, some

Further author information: Corresponding author Ngan T.H. Le. E-mail: thihoanl@andrew.cmu.edu
of them rely on accurate detection of edges and require computing gradient\cite{18,19,11} which might be sensitive to noise. Here we propose to utilize the color information from images by learning a discriminative color subspace, based on which we adopt the minimum average correlation energy (MACE) template matching in frequency domain. Further, we employed an adapted level-set based method which is gradient-free for segmentation. Other than classification, we explored the feasibility of shape analysis based on our segmentation results. Preliminary results show that our proposed framework has potential to detect and segment both normal and abnormal blood cells. The flowchart of our proposed framework is given in Figure 1.

Figure 1. Flowchart of the proposed framework. From a blood smear image, the RBCs and WBCs are first detected in our proposed representation. Our adaptive level set-based method is then apply to WBCs to extract nuclei. In addition to being segmented by the proposed adaptive level set, the RBCs are further studied if it is connected RBCs or abnormal RBCs.

2. METHODOLOGY

2.1 Color representation

Moreover, most state-of-the-art methods have used the pre-defined color spaces such as RGB, simply eliminate the hue and saturation information while retaining the luminance\cite{16,22,25} or HSI (hue, saturation, and intensity) and emphasize on Hue channel\cite{5,7} or Lab color space with focusing on L channel.\cite{17} However, such color spaces are not particularly designed for medical images and they show some weaknesses when displaying WBCs as shown in Figure 2. Hence, finding an appropriate color space for peripheral blood images to present WBCs/RBCs is an importance task in blood cell analysis. Our designed color space (WRB, stands for WBC-RBC-BG) contains three channels corresponding to WBCs, RBCs, and background (BG). The comparison between our proposed color space and other commend pre-defined color like RGB, Lab, HSV is given in Figure 2. In this figure, Figure 2(k) is WBCs color channel, Figure 2(l) is RBCs color channel and 2(m) is BG color channel.

Notice that with representation learning techniques, we can stack standard color space to form new color features in a higher dimensional space which is then reduced to the number of components to be distinguished in the image. In this way, we can fully harness the power of different color spaces and tailor them to specific problem setting of discriminating different components in image by color. One popular choice for subspace learning and dimension reduction is principal component analysis (PCA). Though it is unsupervised and sometimes gives desired color representation, we cannot guarantee that the resulting subspace given by eigenvectors correspond to the first several largest eigenvalues are the most discriminating ones. Here we use linear discriminant analysis (LDA) to learn a discriminative subspace. Other than PCA which learns the direction that maximizes data variation, LDA seeks to project data points onto a direction such that the intra-class variation is maximized and the inter-class variations are minimized at the same time. Mathematically, denote the $i$-th data point or feature vector as $x_i$. We employ a standard two-class Fisher-LDA formulation, where the projection direction $\omega$ is given by maximizing the following objective:

$$J(\omega) = \frac{\omega^T S_B \omega}{\omega^T S_W \omega}$$

(1)

where $S_B$ is known as the between class scatter matrix, whereas $S_W$ is the within class scatter matrix. The definitions of
Figure 2. Comparison between different predefined color spaces and our proposed blood cell color space. 1st row: Original image(a), 2nd row: RGB color space (b-d), 3rd row: Lab color space (e-g), 4th row: HSV color space (h-j), 5th row: our proposed blood cell color space (k-g)

the two scatter matrices are:

\[
S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T \tag{2}
\]

\[
S_W = \sum_{j=1,2} \sum_{i \in C_j} (x_i - \mu_j)(x_i - \mu_j)^T \tag{3}
\]

where \( \mu_j \) is mean of class \( j \). When the image contains more than two components, we instead build one-versus-rest classifiers. The subspace is then formed by stacking these eigenvectors together. Notice that the resulting bases are not imposed to be orthogonal. Figure 2 illustrates the idea of color space learning. It can be seen that in the new color space, WBC nuclei, WBC cytoplasm, RBCs, and background are prominent in corresponding channels.

Notice that one disadvantage of LDA is that it is supervised and thus requires labels of the data. To enable our system to automatically learn an appropriate color space, we first employ Gaussian mixture models (GMM) to estimate the label of each pixel, which is then fed into LDA training. Given the monotonic property of colors in smear images, GMM serve as a powerful and robust classifier. With Gaussian mixtures estimated, the posterior, i.e. the probability of a pixel belongs to a certain component \( C_i \), given the color feature vector \( x \) is compute as:

\[
P(C_i|x) = \frac{P(x|C_i)P(C_i)}{\sum_i P(x|C_i)P(C_i)} \tag{4}
\]

where \( i = \{WBC, RBC, BG\} \), and \( P(C_i) \) are component proportions estimated during GMM iterations. Figure 3 shows the normalized posterior given by GMM.

2.2 Template matching

In 1987 Mahalanobis et al.\textsuperscript{12} developed the minimum average correlation energy (MACE) filter which minimizes the average correlation energy (ACE) of the correlation outputs due to the training images while satisfies the correlation peak constraints at the origin. By minimizing the ACE, it produces a strong peak at the location of a trained object while
everywhere else yield values close to zero. Denote the Fourier transform of \( x_i \) as \( X_i \), the average correlation plane energy is:

\[
E_{av} = \frac{1}{N} \sum_{i=1}^{N} E_i = \frac{1}{N} H^+ \left( \sum_{i=1}^{N} D_i \right) H = \frac{1}{N} H^+ D H,
\]

where the superscript + denotes conjugate transpose, \( n \) is the number of training images from the positive class, vector \( H \) is the Fourier transform of the MACE filter \( h \), and \( D_i = \text{diag}(|X_i|^2) \) (the square of a vector is defined component-wise). More detailed derivation can be found in the original paper, and here we simply give the solution. Basically, the MACE solution seeks to minimize \( E_{av} \) subject to the linear constraint \( X^+ H = u \), where \( u \) is simply set to be a one vector. Finally the resulting MACE filter is as follows in a vector form:

\[
H_{MACE} = D^{-1} X (X^+ D^{-1} X)^{-1} u.
\]

Suppose that each sample is presented in \( R^d \). We perform 2-D FFTs on these images and vectorize them. In Equation 6, \( X \) is a matrix \( d \times N \) whereas \( D \) is a diagonal matrix \( d \times d \) and along its diagonal the average power spectrum of the training images (i.e., average of the magnitude squares of the columns of \( X \)). Column vector \( u \) with \( N \) elements contains the pre-specified correlation peak values of the training images.

At testing time, one common procedure is to compute the peak to sidelobe rate (or PSR). In this paper we simply perform mean-shift on top of the correlation response map after thresholding. Figure 4 shows illustrations of peak responses obtained by MACE filter in our designed color space. The first row of Figure 4 is WBCs peak response whereas the RBCs peak response is given in the second row of Figure 4. For implementation details and how templates are generated, please refer to experiments and results.

2.3 Adaptive level-set segmentation

In our work, we employ a new variational framework developed by M’hiri et al.\cite{13} for detecting the contour. Level set based active contours are superior to other segmentation methods due to the sub-pixel accuracy that they achieve during the fitting process. In this paper, we present an adaptive level set-based segmentation which makes use of the advantages of both global and local active contours and we define the energy function using four terms: (1) a local fitting term, (2) a global fitting term, (3) a contour term, and (4) a regularization term as shown in (7).

\[
F(\phi, f_1, f_2, c_1, c_2) = \alpha F^L(\phi, f_1, f_2) + \beta F^G(\phi, c_1, c_2)
+ \mu \int_{\Omega} \frac{1}{2} |\nabla \phi - 1|^2 \, dx \, dy + \nu \int_{\Omega} \delta(\phi)|\nabla \phi| \, dx \, dy
\]

(7)
In (7), the first term representing the local force of the active contour is expressed using (8), in which \( f_1 \) and \( f_2 \) are the internal and external forces respectively determined by the local information inside and outside the contour. These forces are weighted by \( \lambda_{11} > 0 \) and \( \lambda_{21} > 0 \). Note that gradient operation appears in (7). This is not contradictory to the statement that our model is gradient-free, given that the gradient operates on the level-set function \( \phi \) which is a continuous and smooth function, rather than on the original image \( u_0 \). \( F^L \) is a region-scalable fitting energy function defined by a Gaussian kernel function \( K_\sigma \) defined by \( K_\sigma = \frac{1}{2\pi\sigma^2}e^{-|u|^2/2\sigma^2} \), in which the standard deviation \( \sigma \) can be seen as a scale parameter that controls the region-scalability from small neighborhood to the whole image domain.

\[
F^L(\phi, f_1, f_2) = \lambda_{11} \int_\Omega K_\sigma(f_1 - u_0)^2H(\phi) \, dx \, dy + \lambda_{21} \int_\Omega K_\sigma(f_2 - u_0)^2(1 - H(\phi)) \, dx \, dy \tag{8}
\]

The second term in (7) represents the global force and it is formulated as shown in (9), in which \( c_1 \) and \( c_2 \) represent the average intensities inside and outside the contour respectively and \( \lambda_{12} > 0 \) and \( \lambda_{22} > 0 \) are the parameters that control the force inside and outside the contour.

\[
F^G(\phi, c_1, c_2) = \lambda_{12} \int_\Omega (c_1 - u_0)^2H(\phi) \, dx \, dy + \lambda_{22} \int_\Omega (c_2 - u_0)^2(1 - H(\phi)) \, dx \, dy \tag{9}
\]

To preserve the regularity of the level set function, which is necessary for an accurate computation and a stable evolution of the level set, a regularization term is used as third term in (7). The last term in (7) is used to regularize the zero level set and is thus to derive a smooth contour. \( H \) is the Heaviside function and defined by \( H_\epsilon(x) = \frac{1}{2} \left( 1 + \frac{2}{\pi} \arctan \left( \frac{x}{\epsilon} \right) \right) \). \( \epsilon \) is the Dirac function and defined \( \delta_\epsilon(x) = \frac{1}{\pi} \frac{\epsilon}{\epsilon^2 + x^2} \). To minimize the function \( F(\phi, f_1, f_2, c_1, c_2) \), standard gradient descent is used.

2.4 Shape matching for abnormal blood cell classification

Shape is commonly defined in terms of the set of contours that describe the boundary of an object and considered as a key feature for computer vision applications. Researchers have shown that shape-based or contour-based recognition and matching approaches outperform appearance-based methods in many cases because shape is a strong feature for recognition as psycho-physical studies. Shape matching techniques can be divided into three categories: The first category uses appearance-based methods, their contours described by shape boundaries. The second category measures the similarity between shapes via a part-to-part (or segment-to-segment) matching and junction parameter distribution. The third category captures the part structure by considering the interior of shape boundaries.

In this paper, we choose the inner-distance proposed by Ling and Jacobs\(^{10}\) for shape matching which replaces the Euclidean distance to extend the shape context. Given \( n \) sample points \( P_1, P_2, \ldots, P_n \) on a shape \( P \), the shape context at a point \( P_i \) is defined as a histogram \( h_{P_i} \) of the relative coordinates of the remaining \( n - 1 \) points as shown in (10),

\[
h_{P_i}(k) = \#\{P_j : j \neq i \text{ and } (P_j - P_i) \in bin(k)\} \tag{10}
\]

where bins are uniformly divided log-polar space as shown in Figure 5(a) with the reference point \( P_i \) marked as red star. The histogram \( h_{P_i} \) of point \( P_i \) at the gray bin is the number of points in this area and it is 7 in this example.

Given a connected and closed shape \( P \) and two points \( P_1, P_2 \in P \), the inner distance between \( P_1 \) and \( P_2 \), denoted as \( d(P_1, P_2) \) is defined as the shortest path connecting \( P_1 \) and \( P_2 \) within \( P \). An example of inner distance between \( P_1 \) and \( P_2 \) on the connected and closed shape is given in Figure 5(b).

The shape matching problem is hence formulated as follows: given two shapes \( P \) and \( Q \), their contours described by the points sequences \( P_1, P_2, \ldots, P_n \) and \( Q_1, Q_2, \ldots, Q_m \) respectively. The cost of matching point \( P_i \) on the shape \( P \) to point \( Q_j \) on the shape \( Q \) is defined by the chi-squared distance, as shown in (11).

\[
C(P_i, Q_j) = \frac{1}{2} \sum_{k=1}^{K} \frac{(h_{P_i}(k) - h_{Q_j}(k))^2}{(h_{P_i}(k) + h_{Q_j}(k))} \tag{11}
\]
where $h_{P,i}$ and $h_{Q,j}$ are the shape context histograms of $P_i$ and $Q_j$ respectively. $K$ is the number of histogram bins. The matching from $P$ to $Q$ is done to minimize the total matching cost $\sum_i C(P_i, Q_{\rho(i)})$, in which $\rho(i)$ is a permutation.

With the pair-wise distances computed, theoretically we can then train many different classifiers which are based on distance. In our experiment we utilized k-NN classifier with $k$ equals 1. Another possible direction is to use SVM with Gaussian radial basis function (RBF) kernels, where the $L_2$ distance is replaced by the above defined distance.

3. EXPERIMENTS AND RESULTS

3.1 Cell and nuclei detection

Before MACE template matching is applied to pathology images, training samples are needed to train MACE filters. To extract training samples and perform template matching, we adopt the following procedure: 1. use GMM to obtain a rough segmentation, or a posterior map of how likely a pixel being a cell, on which a global threshold is performed; 2. use simple morphological operations to do some post-processing such as noise removal, as well as connected component analysis to extract well isolated cells; 3. we train MACE filters using obtained samples (one can also enlarge the training samples by image augmentation including resizing, rotating, and shearing images); 4. at test time, we apply mean-shift clustering algorithm to the correlation response map between MACE filter and testing image, after thresholding, and use the centroids returned by mean-shift as detected targets.

We tested our detection algorithm on the publicly available ALL-IDB dataset. We evaluated the cell detection efficiency of MACE template matching using precision and recall, which are defined as follows:

\[
\text{precision} = \frac{\#\{TP\}}{\#\{TP\} + \#\{FP\}},
\]

\[
\text{recall} = \frac{\#\{TP\}}{\#\{TP\} + \#\{FN\}},
\]
Table 1. Precision and recall for RBCs and WBCs detection, evaluated on ALL-IDB

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBCs</td>
<td>98.43</td>
<td>94.99</td>
</tr>
<tr>
<td>WBCs</td>
<td>99.12</td>
<td>99.12</td>
</tr>
</tbody>
</table>

where $TP$, $FP$, $FN$ are true positive, false positive, and false negative, respectively. We evaluated our method on 15 randomly selected images from ALL-IDB dataset. The quantitative results for both RBCs and WBCs are concluded in Table 1.

3.2 Segmentation

We tested our segmentation algorithm on several datasets, including ALL-IDB, WebPath, UPMC, Flicker. Some examples to visualize the performance of the WBCs segmentation on images from different datasets are shown in Figure 7. The F-measure of the proposed WBCs segmentation is 95.8%. Furthermore, the segmentation capacity between Mohamed, et al.’s algorithm and ours is also illustrated in Figure 8. As shown, our method is able to neglect RBCs and segment WBC nuclei more accurately, or in other words, has higher precision and recall rate.

![Figure 7. Some examples of segmentation results from different datasets: 1st col - ALL_IDB1; 2nd col - ALL_IDB2; 3rd col - UPMC; 4th col - WebPath; 5th col - Flicker](image)

![Figure 8. Segmentation performance of Mohamed, et al.'s algorithm (1st row) and our proposed method (2nd row)](image)

4. CONCLUSION

This paper proposes an unsupervised system for quantitative analysis of blood cell images. Although certain algorithms used in the system are supervised as originally proposed, such as LDA and MACE, the labels and examples used in the
training phase are in fact generated from unsupervised methods. Hence the system is overall unsupervised.

There are several advantages of the proposed method. First, we learn a discriminative linear color subspace for better visualization purposes as well as downstream detection and segmentation tasks. Second, we introduced MACE filter in blood cell detection problems. MACE filters produces sharp correlation peaks which is good for localization. Moreover, it trains and operates in frequency domain and thus is computationally efficient. Third, we employed a variational model based on the well-known Chan-Vese\textsuperscript{1} level-set framework. The inherent gradient-free property gives the model robustness over noise. Moreover, as assumed by the model, it serves as a good prior that the given image (in this case, an image patch contains a cell and background) is composed of two parts divided by one continuous closed curve. Nevertheless, limitations exist that some parameters such as thresholds may need to be tuned given different images. We will leave as future work to augment our system with more automated and robust parameter tuning strategies. Furthermore, we plan to improve our shape matching algorithm of cells for having a more reliable and comprehensive analysis of pathology images.

REFERENCES