A RED BLOOD CELL RECOGNITION SYSTEM OF ANEMIA BASED ON BLOOD SMEAR IMAGES

[1]YUNG-KUAN CHAN, [2]MING-JUN SU, [3]CHE-CHIA CHAN

[1,2,3]Nation Chung Hsing University
E-mail: [1]ykchan@nchu.edu.tw, [2]7100029002@mail.nchu.edu.tw, [3]chechia@nchu.edu.tw

Abstract - In blood testing, evaluating whether a person has anemia usually uses hematology analyzer to analysis some factors, such as Hemoglobin (Hb), Red blood cell (RBC), mean corpuscular volume (MCV), etc. However, hematology analyzer is very expensive for general clinics. In order to solve this problem, this thesis proposes a new estimation method based on 2-dimensions blood smear images to develop a mechanism that can help doctors to do diagnosis.The proposed method judges whether the red blood cells belong to one of five anemia blood classes according to the shape of single RBC in blood smear images. These five classes are Acanthocytes, Schistocytes, Circle, Oval, and Teardrop.

Index Terms - Segmentation, RBC, Morphology, Genetic algorithm.

I. INTRODUCTION

Blood testing is easy to be done in large hospitals, but region clinics can not have enough funds to by a hematology analyzer. On the other hand, the inspection method of anemia by blood smear can reduce the workload of inspectors and give some information for doctors. According to the above descriptions, developing this inspection method can reduce the cost of blood testing, money, times, etc. The most important point of this inspection method is to detect the blood smear of anemia. In order to develop this inspection method, it should be known about the kinds of shapes for RBCs in the blood smear at first. According to the shape of RBC, whether the RBCs in blood smear belong to anemia can be known. In this thesis, the shapes of RBCs are classified into five types which are often appeared in the blood smear of anemia. The five types are Acanthocytes-shaped, Schistocytes-shaped, Circle-shaped, Oval-shaped, and Teardrop-shaped. The shape of these five classes is shown in Table 1-1.

Table 1-1. The shape example of five shape classes.

Shape name	Acantho-cytes	Schisto-cytes	Circle	Oval	Teardrop
Shape example					

According to five types of shapes for RBCs, the shapes of RBCs can be used to help doctors to make good diagnoses.

II. THE PROPOSED METHOD

In this section, the proposed method is introduced to segment RBC in the blood smear image. The segmentation method can be divided into three phases: Preprocessing, Single Normal RBC Segmentation, Extraction Light Contact Single RBC. First of all, the results of segmenting RBC regions may be affected by the status of the binary images. To have proper binary image for segmenting RBC regions, the preprocessing is needed as shown in Subsection II.A. After that, the single RBC region is segmented from the preprocessed image and the procedures are shown in Subsection II.B. In order to further segmented some single RBCs which slight contact or overlap with other RBCs, the procedures in Section II.C. are used. After performing three phases, we can use the segmented single RBC to identify what class the single RBC belongs.

A. Preprocessing

In this section, an image I is used to demonstrate all processes. The processes in this phase can be divided into four steps. The first step chooses Green stratifications of RGB color space image for further processing as shown in Subsection A.1). The second step performs Top-hat transform and Bottom-hat transform on G-stratification gray-level image to remove the noise of non-uniform illumination light source ash shown in Subsection A.2). Then, the third step uses Otsu's Thresholding [1] to extract rough RBC regions and the binary I_{TB} gray-level image can be obtained ash shown in Subsection A.3). The final step, the fourth step, uses mathematical morphology[2][3][4] to complete the RBC regions ash shown in Subsection A.4). The details of four steps are shown in the following subsections.

1) Choose the Green Stratification Image of RGB ColorSpace Image

It can be expected that the foreground regions can beclearly separated from the background regions.

2) Combined with Top-hat and Bottom-hat Transform

There usually exists the non-uniform illumination problem for RBC Blood smear images. The proposed method uses Top-hat transform and Bottom-hat transform to solve this problem. The preprocessing step [5] in uses (2-1) to solve the influence problem of non-uniform illumination.

$$I_{TB} = \left(I_G + T_{hat}(I_G)\right) - B_{hat}(I_G) \qquad (2\text{-}1)$$

where I_G is the G-stratification gray-level image, $T_{hat}(I_G)$ denotes as the Top-hat transform result of I_G,

and $B_{hat}(I_G)$ denotes as the Bottom-hat transform result of I_G.

3) Using Otsu's Thresholding

It can bee seen that RBC regions can be distinguished from the blood smear image. In this subsection, Otsu's thresholding is used to extract rough RBC regions and the binary I_{TB} gray-level image can be obtained. To achieve this purpose, histogram distribution in blood smear images are observed and there are always two peaks values existed in the histogram distribution. Under this condition, it is proper to use Otsu's thresholding to classify non-RBC regions and RBC regions in blood smear images. Through the same example above, the threshold T_o obtained by Otsu's thresholding is 154. So, we assumed non-RBC regions as black regions and RBC regions as white regions.

In order to show the importance of performing Top-hat transform and Bottom-hat transform in the second step, the G-stratification gray-level image is gone through Otsu's thresholding directly.

Can be seen that Otsu's thresholding extract rough RBC regions in I_{TB} successfully. The influence of non-uniform illumination is eliminated. However, there are some concave plate regions and further processing will be used in the next step to eliminate this phenomenon.

4) Mathematical Morphology to Let Segmentation Completely

In this subsection, fill holes algorithm is firstly performed on I_{TB} image to fill the concave plate regions of RBC.

Therefore, the following procedure is used to solve this problem:

a) Label the objects in I_{Fill} image. Each RBC region has its unique label number.

b) Select one of labeled numbers to draw a new region containing the RBC region. The new region is denotes as I_{Fill_i} (i=1,2,…,n as the label tag number).

c) Reverse I_{Fill_i} to let pixel values in backgrounds as that in foregrounds are exchange to obtained the new image $I_{Fill_r_i}$.

d) Perform Erosion operator T_E times on image $I_{Fill_r_i}$ by the 3×3 structuring element SE, where the values of all elements in SE are 1. The value of T_E will be set by genetic algorithm[6][7], which will be explained in section III.

e) Calculate the number of object regions in $I_{Fill_r_i}$. If the number is 1, there is no crushing region in I_{Fill_i}, then go to step 2 to handing the next region $I_{Fill_{i+1}}$. Otherwise, go to step 6.

f) Perform Dilation operator T_E times on region $I_{Fill_r_i}$ to obtain some new regions and Combine I_{Fill_i} with those new regions to obtain modified region I_{Fill_i}.

g) Combine all I_{Fill_i}(i=1, 2, … , n as the label tag number) and final image I_{Fill_e} can be obtained.

B. Single Normal RBCs Segmentation

In the field of medicine, the single normal RBC is a three-dimensional pattern which is flat round-shape object with double concave disc. Because the blood smears production method will let one of double concave discs face upward, it can be observed that the edge shape of single normal RBC is likely to be circle in the blood smear images. Therefore, the single normal RBC is segmented from I_{Fill_e}, in this section. There are three steps in this section. The first step finds approximate circle form I_{Fill_e}, . The procedure is described in Subsection B.1). Although the shape of RBC is likely to be a circle , there are some non-RBCs regions whose shape is also likely to be a circle, such as blood platelet regions and noise regions. Therefore, k-means method [8] is used in Subsection B.2) to identify and eliminate them. After that, some RBCs are further eliminated if their regions are smaller than the average area of all RBCs, which is described in Subsection B.3).

1) Finding Approximate Circle

In the blood smear images, it is important to evaluate the average value of single normal RBC. Because the shape of single normal RBC is likely to be circle, (2-2) is used to evaluate the ratio of real area and calculated area where calculation area comes from multiplying π by the square of half of the longest distance. The ratio in (2-2) can be used to check whether the shape of area approximates to a circle.

$$R_{p_i} = A_{o_i}/\pi\left(Lo_{max_i}/2\right)^2 \quad , i = 1,2,\dots,n. \tag{2-2}$$

Where A_{o_i} denoted as the areas of ith RBC region, and Lo_{max_i} denoted as the longest distance of ith RBC regions.

To be more precise about Eq.2-2, let us talk about how to calculate that area of a circle firstly. Its formula is $A_r = \pi r^2$ where r is the radius of the circle. Now, we replace r as half of the longest distance, that is $\left(Lo_{max_i}/2\right)$, and the calculated area must be larger or equal to Ar. The equal condition occurs when the shape of the area is circle. Therefore, if the value of R_{p_i} is close to 1 , the shape of area approximates to a circle. In this thesis, we set a threshold T_r to distinguish circle shape and non-circle shape as shown in (2-3). The value of T_r will be set by genetic algorithm, and it will be explained in section III.

$$\begin{cases} R_{p_i} \geq T_r & , \text{thinks as round.} \\ R_{p_i} < T_r & , \text{thinks as non} - \text{round.} \end{cases} \tag{2-3}$$

2) Classify RBC regions by K-means Finding Approximate Circle

In this subsection, the concept of k-means clustering is used to classify the types of objects in I_{round}. First of all, Lo_{max_i} shown in (2-2) is calculated as a feature to classify as three types of objects, including blood cells (BC), blood platelets (BP), and noises. According to the observation, Lo_{max_i} of BC is the bigger than that of BP and noises in field of medicine. Therefore, we use the following steps to classify the types of objects in I_{round}:

a) Sort values of all Lo_{max_i} , $i = 1,2,\dots,n$ in a descending order.

b) Set the initial center of BC class as the same longest

distance of the element at the front quarter, that is $C_{BC_0} = Lo_{max_{\frac{1}{4}n}}$. Set the initial center of BP class as the longest distance of the element at the medium, that is $C_{BP_0} = Lo_{max_{\frac{2}{4}n}}$. Set the initial center of noise class as the same longest distance of the element at the last quarter, that is $C_{noise_0} = max_{\frac{3}{4}n}$.

c) Compute the distance between Lo_{max_i} with all three classes. Assign Lo_{max_i} to the class where Lo_{max_i} has the minimum distance. Therefore, three classes are BC class $X_{BC_j}, j = 1,2,\dots,k_1$, BP class $X_{BP_j}, j = 1,2,\dots,k_2$, and noise class $X_{noise_j}, j = 1,2,\dots,k_3$. Moreover, $k_1 + k_2 + k_3 = n$.

d) Compute the new centers for three classes by calculating its mean value, $C_{BC_r} = avg\left(X_{BC_j}\right), j = 1,2,\dots,k_1$, $C_{BP_r} = avg\left(X_{BP_j}\right), j = 1,2,\dots,k_2$, and $C_{noise_r} = avg\left(X_{noise_j}\right), j = 1,2,\dots,k_3$.

e) Repeat step 3 and step 4 until $C_{BC_r} = C_{BC_{r-1}}$, $C_{BP_r} = C_{BP_{r-1}}$, and $C_{noise_r} = C_{noise_{r-1}}$ or the repeating time reaches the predetermined times r.

After performing the above steps, the centers of three classes can be known and each object can be checked which class it belongs to. Therefore, the non-RBC object can be eliminated. However, it still has a problem for the ease when the centers of any two classes are closed. In this case, two closed classes are merged into one class. More precisely, between any two values of center classes, if the value which comes from multiplying the smaller value by K_c is larger than the larger value, two classes are merged. In the aspect of value of K_c, its value will be determined by genetic algorithm in section III.

3) Eliminate Smaller Regions than the Average Area of RBC Regions

To eliminate RBCs with smaller area, the mean values of RBC regions and mean length of radius should be calculated firstly. The mean value M_{round} of objects comes from computing all the pixels in them and the mean area length of radius comes from computing $R_{round} = Lo_{max_i}/2, i = 1,2,\dots,n_{round'}$. Then, if the region areas are small than the value which comes from multiplying K_l by the mean of single normal RBC, those areas are eliminated. In the aspect of value of K_l, its value will be determined by genetic algorithm in section III.

C. Segmentation Light Contact Single RBCs

After obtaining image I_{Fille_r}, some single RBCs with slight contact or overlap each other should be segmented, some RBCs with seriously overlap should be removed, and the nuclear of WBCs should also be removed. In this section, three steps are proposed to solve the above problem. This section can be divided into as following: New Adaptive Opening Algorithm, Conditional Merger Algorithm, and Eliminate Larger Object Regions and Nuclear of WBCs Regions. The procedures of segmenting multi-RBC regions are introduced in Subsection C.1), which is called New Adaptive Opening Algorithm. After through procedures in Subsection C.1), it maybe segments the multi-RBC regions too much. Therefore, we use procedures in Subsection C.2) to merge the smaller regions into one region, which is called Conditional Merger Algorithm. Then, un-single RBCs regions and the nuclear of WBCs regions are removed as described in Subsection C.3). This step is called Eliminate Larger Object Regions and Nuclear of WBCs Regions.

1) New Adaptive Opening Algorithm

a) Label object regions in I_{Fille_r} as the label ID $tag_i, i = 1,2,\dots,n$. Determine erosion times $E_r = R_{round} \times R_k$ by a parameter R_k which is generated by genetic algorithm.

b) Perform erosion once at a-th object and label the remaining objects again as new label ID $Tag_j = tag_a + (j-1), j = 1,2,\dots,m$, where m denoted as the number of remaining objects. If m larger than 1, go to step A. Otherwise, go to step B.

(1) Find the edges of each remaining object Tag_j. Compute the distance between the pixels in eroded object regions and edges of each remaining object. Assign the pixels in eroded object regions as the Tag_j ID if the pixel has minimum distance with edges of the remaining object with Tag_j ID. For object with tag_i ID larger than tag_a ID, we reset its tag_i ID as $tag_i = tag_a + (m-1), i = a+1, a+2,\dots,n$.

(2) Find eroded object regions labeled tag ID as remaining object Tag_j.

c) If object regions area still larger than the value which comes from multiplying K_h by the mean of single normal RBCs, go to step 2 until the erosion times is equal to E_r.

d) Repeat all the object regions $tag_i, i = 1,2,\dots,n$ in I_{Fille_r} by step 2 and 3 to generate a label map Map_l. The label map Map_l is an image with labeling tags.

2) Conditional Merger Algorithm

In previous subsection, we can see there are some objects with smaller regions than the predetermined area which comes from multiply K_l by the mean of single normal RBCs. In this subsection, we proposed a conditional merging algorithm as follows to solve this problem:

a) Find all the objects with smaller regions than the predetermined in the label map of I_{Fille_r}.

b) Record all tag IDs (including background label tag ID 0) and their corresponding pixel amounts of the regions which are next to the small region in the label map of I_{Fille_r}.

c) Calculating the mode of label tag ID s in step 2, and set all the smaller regions pixel's label tag ID as the mode's label tag ID

d) Re-arranged the label tag ID after above steps to generated the new label map Map_m

3) Eliminate Larger Object Regions and Nuclear of

WBCs Regions

a) After the above procedures, some RBCs with seriously overlap should be removed, and the nuclear of WBCs should also be eliminated. First of all, we eliminate the some objects with larger regions than the predetermined area which comes from multiply K_l by the mean of single normal RBCs. Moreover, the objects whose regions touch to the image border are also eliminated. In order to eliminate nuclear of WBCs regions, we set a threshold $T_{RBC_g} = mean(I) - H_g \times std(I)$ of RBCs regions gray-level value to judge whether the regions is nuclear of WBCs regions or not. If there are above P_g% gray-level pixels whose pixel values are smaller than T_{RBC_g} in a region, the region is treated as nuclear of WBCs regions and should be eliminated.

D. Concave Plate Segmentation

When obtaining the single RBCs label map from previous section, we can use it as a map to find the concave plate in it. According to the object regions, we can use contrast enhancement formula [9] (2-4) in the object ranges to enhancement the concave plate regions:

$$I_e^{'} = [(I_e - min\ (I_e))/(Max\ (I_e) - min\ (I_e))]^{r_e} \times 255, \quad (2\text{-}4)$$

where r_e is the parameter generated by genetic algorithm, I_e means the pixel that need to enhance contrast images, and $Max\ (I_e)$ and $min\ (I_e)$ denote the largest and smallest gray-level value of I_e.

III. RECOGNITION

In the section III, we uses Map_s segmentation results to classify each single RBC to its type. According to the shape of RBCs in blood smear, the types of recognition classification in this thesis have been divided into five, including Acanthocytes-shaped, Schistocytes-shaped, Circle-shaped, Oval-shaped, and Teardrop-shaped. The shapes of these five types are shown in Table 1-1.

To classify the types of RBCs, the features of RBCs should be extracted first in Subsection III.A. After combining those features by using the weights which was by genetic algorithm, the types of RBCs can be determined. The details are described in Subsection III.B.

A. Features Classification

In order to classify the shape of RBCs, we proposed 11 features to identify region patterns of RBC. Taking a RBC object as shown in Fig.3-1 as an example, P_e denotes as perimeter of the RBC regions and Lo_{max} denotes as the max distance of the RBC regions. RBC MBR represents the minimum bounding rectangle of RBC.

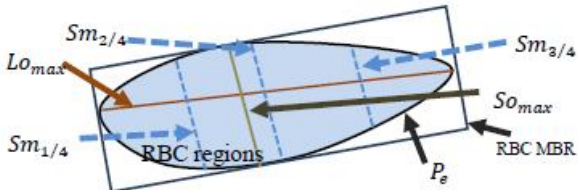

Figure.3-1. The example of RBC.

The symbols defined in Fig.3-1 are used to describe the 11 features as following:

Feature 1.In order to judge whether the shape of RBC region is a circle or not, we use (3-1) to evaluation and A_{RBC}denote the area of RBC region. In the theory, F_1 value is in the range from 0 to 1, and if the shape of RBC region is similar to a circle, the value in Eq. 3-1 will be close to 1.

$$F_1 = A_{RBC}/(\pi(Lo_{max}/2)^2) \quad (3\text{-}1)$$

Feature 2.In order to judge whether the edge of RBC region is a multilateral bulge shape or not, we use (3-2) to evaluation. If the distance between each edge pixel to the center of circle is the same, ${P_e}^2/4\pi$ value is close to A_{RBC}, that is, F_2 value is close to 1.

$$F_2 = A_{RBC}/({P_e}^2/4\pi) \quad (3\text{-}2)$$

Feature 3.In order to judge whether the shape of RBC is a long strip, we set the large side of the RBC MBR as the major axis of RBC and the shorter side of the RBC MBR as the minor axis of RBC as shown in Fig.3-1. (3-3) is used to calculate the ratio of major axis and minor axis. If value of F_3 is close to 1, the shape of RBC is a circle. Otherwise, the shape of RBC may be a long strip.

$$F_3 = So_{max}/Lo_{max} \quad (3\text{-}3)$$

Feature 4. In the Fig.3-1, the major axis of RBC is divided into four parts. The three vertical lines which are perpendicular to major axis at dividing positions can be found. The lengths of three vertical lines denoted as $Sm_{i/4}$, $i = 1,2,3$. The fourth feature comes from dividing the minimum value by Lo_{max_i} as shown in (3-4).

$$F_4 = min\left(ASC\left(Sm_{i/4}/So_{max}\right)\right) \ , i = 1,2,3 \quad (3\text{-}4)$$

Feature 5. The Feature 5 comes from dividing the second minimum value of vertical lines by Lo_{max_i} as shown in (3-5).

$$F_5 = second_min\left(ASC\left(Sm_{i/4}/So_{max}\right)\right) \ , i = 1,2,3 \quad (3\text{-}5)$$

Feature 6. The Feature6 comes from dividing the second maximum value of the vertical line by Lo_{max_i} as shown in (3-6).

$$F_6 = Max\left(ASC\left(Sm_{i/4}/So_{max}\right)\right) \ , i = 1,2,3 \quad (3\text{-}6)$$

Feature 7. In the Fig.3-4, RBC MBR area is divided into four parts. They are denoted as $Oba_{i/4}$, $i = 1,2,3,4$. The Feature7 comes from dividing the minimum value of four parts by the quarter of RBC MBR area as shown in (3-7).

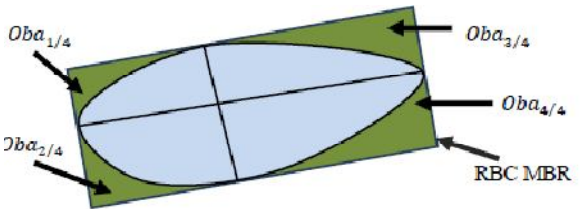

Figure.3-4. The 1/4, 2/4, 3/4, 4/4 minimum rectangle bounders outer area.

$$F_7 = min\left(ASC\left(Oba_{i/4}/Ob_{1/4area}\right)\right) \quad , i = 1,2,3,4 \tag{3-7}$$

Feature 8. The Feature 8 comes from dividing the second minimum value of four parts by the quarter of RBC MBR area as shown in (3-8).

$$F_8 = second_{min\left(ASC\left(Oba_{i/4}/Ob_{1/4area}\right)\right)}, i = 1,2,3,4 \tag{3-8}$$

Feature 9. The Feature 9 comes from dividing the second maximum value of four parts by the quarter of RBC MBR area as shown in (3-9).

$$F_9 = second_{Max\left(ASC\left(Oba_{i/4}/Ob_{1/4area}\right)\right)}, i = 1,2,3,4 \tag{3-9}$$

Feature 10 The Feature 10 comes from dividing the maximum value of four parts by the quarter of RBC MBR area as shown in (3-10).

$$F_{10} = Max\left(ASC\left(Oba_{i/4}/Ob_{1/4area}\right)\right), i = 1,2,3,4 \tag{3-10}$$

Feature 11. In the theory, the convex hull [10] region is larger than RBC regions. As the formula shown in (3-11), we can use it as the feature to classify RBC.

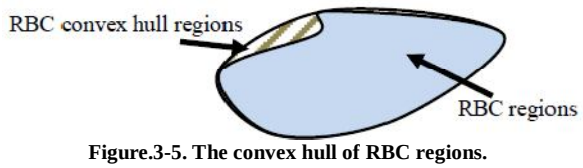

Figure.3-5. The convex hull of RBC regions.

$$F_{11} = A_{RBC}/A_{convex_hull} \tag{3-11}$$

After generating the above features, some data are used to train so as to classify them into five types of RBC. The numbers of Training data are shown in Table 3-2.

Table 3-2. The training amounts of five classes.

Shape name	Acantho -cytes	Schisto -cytes	Circle	Oval	Teardrop
Training amounts	10	10	10	10	10

The distance of RBCs in different two images through those features distance can be calculated by (3-12).

$$minD_{F_k} = W_i \times \left|F_{ij} - \overline{F_{tk}}\right|^{r_i} ,$$
$$i = 1,2,\dots,11 , j = 1,2,\dots,n, k = 1,2,\dots,5 \tag{3-12}$$

where W_i denotes the weight of each feature distance, r_i denotes the gamma index of each feature distance, F_{ij} denotes the jth feature value of training data in ith class, and $\overline{F_i}$ denotes the feature mean value of ith class. The values of W_i and r_i are trained by genetic algorithm.

After getting weights from training these data, (3-12) is used to calculate the type of the testing data in blood smear images.

IV. THE EXPERIMENTAL RESULTS

The experimental results include two parts: Single RBC Segmentation Results and Shape of RBC Classification Results. The evaluation criteria include True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) .

A. Single RBCs Segmentation Result

In the first experiment, the parameters for segmentation are trained by genetic algorithm as shown in section II. The training data are five blood smear images. The parameters determined by the genetic algorithm are $T_E = 4$, $T_r = 0.85$, $K_c = 1.3$, $K_l = 0.65$, $K_h = 1.5$, $R_k = 0.45$, $r_e = 3$, $H_g = 2$, and $P_g = 57$. And we use other blood smear images to test whether those values are good for segmentation.

The proposed segmentation method is evaluated by comparing the results of the proposed segmentation method and ground truth images which are manually marked by doctors. The performance evaluation contains four statistical criteria, namely, Misclassification Error (ME) [11] , Relative Area Error (RAE) [11] , Relative Distance Error (RDE) [12], Modified Hausdorff Distance (MHD) , which can be calculated using (4-1)~(4-4) respectively.

$$ME = 1 - ((TN + TP)/(TN + FN + FP + TP)) \tag{4-1}$$

$$\begin{cases} RAE = (FP - FN)/(TP + FP) & , if\ FP \geq FN \\ RAE = (FN - FP)/(TP + FN) & , if\ FP < FN \end{cases} \tag{4-2}$$

The symbol TP is defined as the pixels of RBC regions in the ground truth and our method also determine them as the pixels of RBC regions. The symbol FP is defined as the pixels of RBC regions in the ground truth but our method determine them as the pixels of background regions. The symbol TN is defined as the pixels of background regions in the ground truth but our method determine them as the pixels of RBC regions. The symbol FN is defined as the pixels of background regions in the ground truth and our method also determine those pixels as the pixels of background regions.

$$RDE = \left(\left(\left(\sqrt{\sum_{f_G \in F_G} min_{f_P \in F_P} \|f_G - f_P\|^2}\right)/(|F_G|)\right) + \left(\left(\sqrt{\sum_{f_G \in F_G} min_{f_P \in F_P} \|f_P - f_G\|^2}\right)/(|F_P|)\right)\right)/2 \tag{4-3}$$

$$MHD = \frac{1}{|F_G|}\sum_{f_G \in F_G} max_{f_G \in F_G}\, min_{f_P \in F_P} \|f_G - f_P\| \tag{4-4}$$

where $|F_G|$, $|F_P|$ is defined as the pixel numbers of ground truth and proposed segmentation method, f_G, f_P is defined as the coordinate pixel in ground truth and proposed segmentation method coordinate pixel, F_G, F_P is defined as the pixel sets of ground truth and proposed segmentation method.

The smaller the four values have, the better our segmentation method get. In our experimental results, the means of ME is 14.56%, the means of RAE is 6.05%, the means of RDE is 4.469, and the means of MHD is 1.727 in our segmentation method. And five experimental results of blood smear image by our segmentation method are shown in Table. 4-1~4-2.

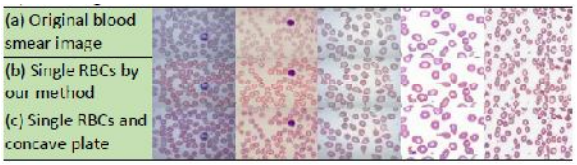

Table.4-1.The experimental results of blood smear image by our segmentation method.

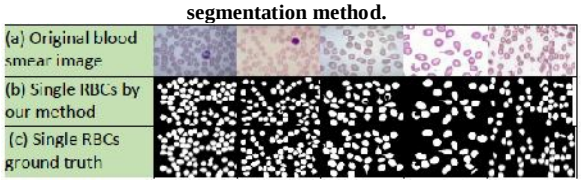

Table.4-2. Compared with our segmentation results and ground truth.

B. Shape of RBC Classification Result

In the second experiment, the parameters for classification are trained by genetic algorithm as shown in section III. The parameters determined by the genetic algorithm are W_i , $i = 1,2,\ldots,11$ and r_i , $i = 1,2,\ldots,11$. They are shown in Table 4-3.

W_i	0.0847457	0.0508474	0.0677966	0.1016949	0.0847457	0.1016949
r_i	0.1612903	0.0483870	0.0806451	0.1129032	0.0322580	0.0967741
	7	8	9	10	11	
W_i	0.0847457	0.1016949	0.0847457	0.1186440	0.1186440	
r_i	0.1290322	0.0806451	0.0483870	0.1290322	0.0806451	

Table 4-3. The classification parameters in our method.

The proposed segmentation method is evaluated by comparing the results of the proposed segmentation method and ground truth images which are manually marked by doctors. The performance evaluation contains three statistical criteria, namely, sensitivity [13][14], specificity [13][14] and F-value [13][14], which can be calculated using (4-5)~(4-6), respectively.

$$\text{Sensitivity} = (TP/(TP + FN)) \tag{4-5}$$

$$\text{Specificity} = (TN/(TN + FP)) \tag{4-6}$$

$$F = 2 \times \frac{\text{Sensitivity} \times \text{Specificity}}{\text{Sensitivity} + \text{Specificity}} \tag{4-7}$$

We compare each blood smear image with their classification ground truth images to measured sensitivity, specificity, and F-value. According to it should evaluate the effectiveness of our classification method, we divided the classification result into two cases: case-1 is using our segmentation result, case-2 is using our segmentation result which removing the wrong part of segmentation. In our experimental results, the mean of sensitivity is 94.90%, specificity is 83.84%, F-value is 84.81% in case-1 classification method; the mean of sensitivity is 96.07%, specificity is 92.23%, F-value is 93.05% in case-2 classification method. The sensitivity, specificity, and F-value of each shaped classes' showing as Table 4-4 and Table 4-5.

	Acantho -cytes	Schisto -cytes	Circle	Oval	Teardr -op	means
Sensitivity	90.57%	91.96%	97.16%	73.06%	91.97%	94.90%
Specificity	99.03%	100.00%	81.43%	98.70%	99.21%	83.84%
F-value	90.06%	93.39%	84.85%	77.94%	94.83%	84.81%

Table 4-4. The classification result by case-1

	Acantho -cytes	Schisto -cytes	Circle	Oval	Teardr -op	means
Sensitivity	91.67%	95.00%	97.54%	82.83%	96.71%	96.07%
Specificity	99.05%	99.77%	90.78%	99.10%	99.29%	92.23%
F-value	92.00%	96.59%	93.38%	86.80%	97.77%	93.05%

Table 4-5. The classification result by case-2

CONCLUSION

In our experimental results, the means of ME is 14.56%, the means of RAE is 6.05%, the means of RDE is 4.469, and the means of MHD is 1.727 in our segmentation method. Moreover, the mean of sensitivity is 94.90%, specificity is 83.84%, F-value is 84.81% in case-1 classification method; the mean of sensitivity is 96.07%, specificity is 92.23%, F-value is 93.05% in case-2 classification method. The experimental results show that our method is effective in segmenting the regions of single RBC and classifying this segmentation results into the five shape types of single RBC successfully. By the automatic classifying the shape of RBCs in the blood smear images, our method reduces the manpower demand of classifying the shape of RBCs in the blood smear. In this way, our proposed method can help doctors to do their diagnosis based on the blood shapes of single RBC. Through the experimental results, it can be seen that the inspection method of anemia by using blood smear image can be done in future.

REFERENCES

[1] N. Otsu, "A Threshold Selection Method from Gray-Level Histogram," IEEE Transactions on System Man Cybernetics, vol. SMC-9, no. 1, pp.62-66, 1979.

[2] J. Serra, "Image Analysis and Mathematical Morphology", 1982.

[3] E. R. Dougherty , "An Introduction to Morphological Image Processing" , 1992.

[4] G. J. F. Banon, J. Barrera, U. M. Braga-Neto, "Mathematical Morphology and Its Applications to Signal and Image Processing", 8th International Symposium on Mathematical Morphology (ISMM'07), 2007.

[5] M. D. Saleh, C. Eswaran, "An Automated Decision-support System for Non-proliferative Diabetic Retinopathy Disease

Based on MAs and HAs Detection", Computer Methods and Programs in Biomedicine, vol. 108, no. 1, pp. 186-196, 2012 .
[6] M. Melanie, "An Introduction to Genetic Algorithms" Cambridge, MA: MIT Press, 1996.
[7] W. Darrell, "A Genetic Algorithm Tutorial" Statistics and Computing, vol.4, no. 2, pp. 65–85, 1994.
[8] J. B. MacQueen, "Some Methods for Classification and Analysis of Multivariate Observations" Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability 1. University of California Press, pp. 281–297, 1967.
[9] Y. K. Chan, P. Y. Pai, C. C. Liu, Y. S. Wang, C.W. Li, L.Y. Wang, "Fluorescence Microscopic Image Cell Segmentation," International Journal of Future Computer and Communication, vol. 1, no. 1, pp.72-75, 2012.
[10] Andrew, A. M., "Another Efficient Algorithm for Convex Hulls in Two Dimensions", Information Processing Letters, vol. 9, no. 5, pp. 216-219, 1979.
[11] M. Sezgin, and B. Sankur, "Survey over Image Thresholding Techniques and Quantitative Peformance Evaluation," Journal of Electronic Imaging, vol. 13, no. 1, pp. 146-165, 2004.
[12] P. Yan, X. Zhon, M. Shah, and T. C. Wong, "Automatic Segmentation of High-Throughput RNAi Fluorescent Cellular Images,'' IEEE Transcations on Information Technology in Biomedicine, vol. 12, no. 1, pp. 109-117, 2008.
[13] L. Costaridou, "Medical Image Analysis Methods The Electrical Engineering and Applied Signal Processing Series," CRC Press, 2005.
[14] T. Athanasiou, A. Darzi, "Evidence Synthesis in Healthcare: A Practical Handbook for Clinicians," Springer, 2011.

★★★