MATEC Web of Conferences **44**, 01086 (2016)
DOI: 10.1051/matecconf/20164401086

Image processing and feature extraction of microscopic

Chun Hua Shi

Langfang Health Vocational College, Hebei, China, Email:545682864@qq.com

Abstract. This paper analyzes the characteristics of white blood cells, diseased cells during feature extraction, the introduction of anti-corrosion factor for microscopic image processing, experiments show that the method of classification and extraction of diseased cells have better results.

1 Introduction

In order to identify blood cells, and analyzed and classified, it is necessary to find the blood cells can be distinguished from the image feature, in order to describe the blood cells, a process called feature extraction. Reflect the characteristics of a lot of blood cell images, in addition to blood cells can be extracted directly from the segmentation feature image, but also can get more specific blood cells characterized by transformation. Commonly used in blood cell analysis and classification features are morphological features, especially the optical density, color, texture characteristics.

Because the computer automatically detect leukemia has been a problem of image processing and pattern recognition, although there are some automated blood cell morphology analysis system, but the accuracy of its analysis is not very high. We hope that from the simplest and most easily recognizable type of white blood cells begin to analyze.

2 Effective features leukemia detected

Characteristics described in the previous section is used to identify the type of white blood cell, not all effective. In this trial, considering three types of cells, segmented neutrophils granulocyte cell line, rod-shaped blood granulocytes and in dealing with the film being damaged cells.

In the study of these three leukocyte classification problems, if we do not use the optical density and texture information, only it uses geometric features. But in the feature geometry, since the cytoplasm is not much impact on the sub-classification, some features are not available. Through repeated comparison test, we finally chose the four characteristics: Area nucleus, roundness nucleus, representing the number of nuclei nucleus and nucleus of convex ratio.

The nucleus area S(1)

$$S(1)=\sum_{(i,j)\in R_c} f(i,j),\quad \text{其中} f(i,j)=\begin{cases}1 & (i,j)\in R_c\\ 0 & (i,j)\notin R_c\end{cases} \tag{2-1}$$

Formula Rc is nucleated blood cell nucleus region. S (1) is the sum of all the number of pixels Rc, which reflect the different cell nuclei size difference in the two-dimensional plane.

2. Roundness nucleus S(2)

$$S(2)=P^2/S(1) \tag{2-2}$$

P is the perimeter of the nucleus in the formula, which is calculated as

$$P=N_e+\sqrt{2}N_o \tag{2-3}$$

Formula E is the sum of the even number of nuclei area boundary chain code, No digital sum is an odd number.

Roundness is close to 1, the region is approximately circular, the greater the roundness otherwise explain the irregular shape of the nucleus. It is to a certain extent, describes the compactness of the nuclear region.

3. Nuclear accounted convex S(3):

$$S(3)=S(1)/Co(S(1)) \tag{2-4}$$

Formula Co (S (1)) to be able to just contain the nucleus area of a circle. Solving the circle area, first find the center position nucleus (that is the nucleus of the center of gravity), in the center and gradually increase the radius of the circle, the circle can have exactly know where to find the nucleus, and the nucleus is tangent to the circle.

Accounting projections reflect the depressed level than the nucleus. The higher the value of S (3), indicating the extent of the smaller nuclei depression. In general, the smaller the cells the more naive depressed, the greater the value; otherwise the value is smaller. In our experiment, the cells and the degree of nuclear recessed rod-shaped leaf granulocyte granulocytes are large and small degree of damaged cells nucleus depression.

The number of nucleus nucleus S(4):

[a] yuanquanbo@126.com

$$S(4)=\sum_i N(i) \quad \text{其中} N(i)=\begin{cases}1 & NR_i \in R_c \\ 0 & NR_i \notin R_c\end{cases}$$

(2-5)

Formula NRi different cell nucleus area under one area, S (4) reflects the number of different cell nucleus area of the same region, which is divided leaf cells with other cells is an important feature of the difference between core granules.

3 Effective features access

3.1. Analyzing nucleus affiliations - merge and cut the cell nucleus

In order to make the computer to find a number of small cells from a large figure, and acquires the cell characteristics, we must find out with some policy contains several images in a blood cell, and determine their location. But contains neutrophils and segmented neutrophils rod-shaped, even relatively large degree of aggregation of blood cells in the image, the resulting image after segmentation are often found in more than one nucleus together, which belong to several nuclear zone several cells. Several nuclear area which belong to the same cell, which constitutes a problem. While the human eye which for us is very easy to judge things, but by the computer to recognize, you need a certain strategy.

As illustrated, one image is included in the two cells 3-1, wherein a cell nucleus region composed of three, and the other consists of a single cell nucleus region.

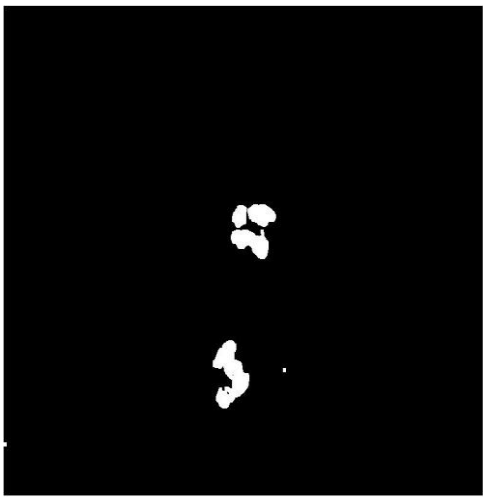


Figure 3-1 before the merger Images

Merger process:

1. image mark merge before the connected regions 1, ..., n.

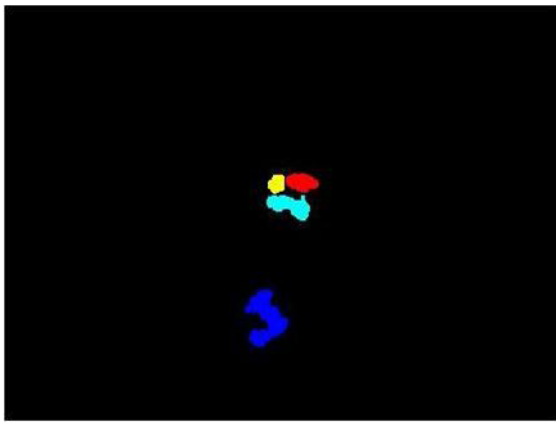
2. Calculate the size of each connected region A1, ..., An.

3. For the communication area of less than a certain value of As, considered to be noise, omitted to consider (As represents the area of the noise threshold point).

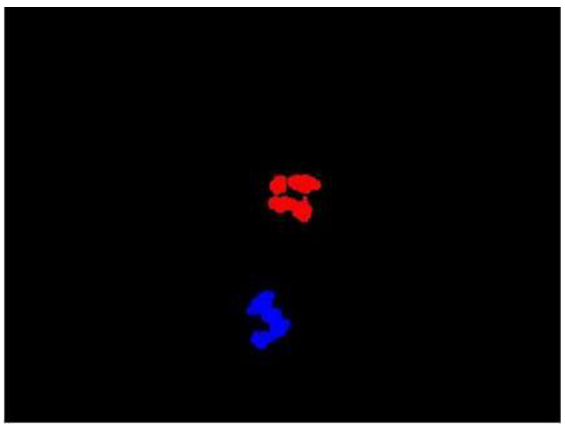
4. For each of the remaining regional connectivity i, if the area is less than a certain value inf (Ac) gradually outward to search along the outer boundary of the communication area. If a limited number of steps, find another connected region j, and Aj in this area together with the communication area Ai area itself does not exceed sup (Ac), and the two connected regions circumscribed circle of diameter not exceeding D, on own mark to j; if not find another one connected region in a limited number of steps, and to give up the search. (Inf (Ac) as a nucleus area of the lower bound, sup (Ac) is the upper bound nucleus area, D is the diameter of the circumscribed circle of the upper bound nuclei.)

5. The cycle above search process until each connected region have been treated once.

6. The re-tagged images, the cells were cut to give each cell contains only a small map.



(a) Mark image pre-merger



(b) Mark images merged

Figure 3-2 is an image processing of Figure 3-1, in which Fig (a) is the result of Step 1 above labeled, and Fig. (B) is the result of re-marked. From the figure we can see that each cell in addition to being a very well-marked, can also be found in Figure 4-1 small particles are also filtered out the noise.

The image after the above steps, will give a few of each cell contains only a small diagram. Small figure characteristics of each cell extract.

3.2. Feature Extraction and problems of encountered

The article mentions four kinds of characteristics in the use of this test: Area nucleus, roundness nucleus, the nucleus and the nucleus than the number of accounts convex nuclear area. Leukocytes has been cut out small image, each graph contains only one white blood cells, so it is easy to calculate its nuclear area, perimeter nucleus, and the circumscribed circle, then, the first three features can be easily extracted.

But the number of nucleus nucleus fourth feature is not capable of very accurate extraction. Many literature [1] mentioned to use this feature to do split, but did not specify how to obtain this feature.

4 Introduction of new features - corrosion factor

Since the number of nuclei nucleus area is difficult to accurately extract the feature, but we try to extract the feature process found that although the number of nucleus can not be accurately calculated, but we found that if the nucleus in multiple etching process very characteristic changes.

For example, leaf neutrophils after 1-5 times corrosion may be disconnected; the rod neutrophils may disconnect after 3-7 times the corrosion, but may continue to open; damaged cells (the basic nucleus on spherical) through a limited number of corrosion will not be disconnected. We call this feature called anti-corrosion properties.

Calculation against corrosion characteristics are: checking cell nucleus now the number, if the number of nuclear area greater than or equal to 2, anti-corrosion factor is defined as the number of current nuclear area, otherwise start the following procedures.

Etching nuclei judge has several regional connectivity. After the 1st of corrosion, which communicates region into two, so it will be an anti-corrosion factor as 2-k, and jump out; if communication is still a number of regions, the corrosion factor is 1-k, and continue corrosion; after the i-th corrosion, if the number of regional connectivity to 2, then the corrosion factor becomes 2-ki, if the number of connected region is still 1, then the corrosion factor is the current value minus ki. (K is the penalty factor, select their own.)

We use the above method of leukocyte nuclei were analyzed to obtain the results of the analysis of: segmented neutrophils corrosion resistance characteristics substantially between 1-2, rod-shaped granulocytes corrosion resistance between 0-1, and corrosion damaged cells are basically below zero. Table 4-1 shows the characteristics of leukocytes 30, wherein the first to 10 for the leaf neutrophils, mark it as a Class 1; Article 11 and Article 20 of the rod neutrophils, which is labeled 2 category; first 21-30 of damaged cells, which marked three categories.

		area	Roundness	Accounting for more than projections
1	Segmented neutrophils	3316	36.103	0.57281
2		2980	31.834	0.41137
3		2431	38.016	0.5089
4		1875	34.68	0.55001
5		3571	33.913	0.48094
6		3243	34.399	0.40665
7		2401	40.284	0.46218
8		3047	31.947	0.62708
9		2527	42.315	0.45912
10		3765	54.987	0.51315
11	Rod-shaped granulocytes	2451	27.369	0.44981
12		2699	37.94	0.52054
13		2426	26.176	0.53756
14		2926	18.395	0.53896
15		3105	24.89	0.69092
16		2453	40.708	0.5541
17		2781	19.354	0.50518
18		3129	23.645	0.62269
19		2744	17.799	0.4749
20		2763	30.859	0.5784
21	Damaged cells	2761	11.997	0.86039
22		3247	16.577	0.52363
23		2738	10.805	0.81952
24		3878	15.988	0.86756
25		4247	13.789	0.69886
26		5288	14.932	0.65592
27		3442	12.329	0.84925
28		4362	13.315	0.60066
29		2476	12.653	0.7668
30		2575	12.304	0.76455

Corrosion resistance	Actual category	Predicted category
2	1	1
2	1	1
2	1	1
2	1	1
1.25	1	2
2	1	1
2	1	1
2	1	1
2	1	1
1.5	1	1
1	2	2

0.75	2	2
1.25	2	2
0	2	2
0.5	2	3
1	2	2
-0.25	2	2
0.75	2	2
-0.5	2	2
0.5	2	2
-1.75	3	3
-1.75	3	3
-1.75	3	3
-1.75	3	3
-1.75	3	3
-1.75	3	3
-1.75	3	3
-1.75	3	3
-1.75	3	3
-1.75	3	3

Table 4-1 Characteristics and Classification of three cells

5 Summary

We see that the nucleus of corrosion resistance between the three cells have different values, for classification, will play a good role.

References

[1] Wang Haojun Blood Cells automatic identification and classification methods. Xi'an Jiaotong University doctoral thesis, 2002.
[2] G. H. Landerweerd, B. S. Gelsema, J. F. Brenner, et al. Pattern recognition of nucleated cells from the peripheral blood. Pattern Recognition, 1983, 16(2): 131-140.
[3] D. Seigneurin, C. Gauvain, G. Brugal. A semi-quantitative analysis of human bone marrow granulocytic linage using the SAMBA 200 cell image processor: I. The normal maturation sequence. Analyt Quant Cytol, 1984, 6(3): 168-178.
[4] M. Beksac, V. B. Tipi et al. An artificial intelligent diagnostic system on differential recognition of Hematopoietic cells from microscopic images. Cytometry, 1997, 30: 145-150.
[5] P. E. Pavlova, K. P. Cyrrilov, I. N. Moumdjiev. Application of HSV colour system in identification by colour of biological objects on the basis of microscopic images. Computerized medical imaging and Graphics, 1996, 20(5): 357-364.

1. Luigi T.De Luca, *Propulsion physics* (EDP Sciences, Les Ulis, 2009)
2. F. De Lillo, F. Cecconi, G. Lacorata, A. Vulpiani, EPL, **84** (2008)