Automated Identification and Classification of White Blood Cells (Leukocytes) in Digital Microscopic Images

P.S.Hiremath
Dept. of Computer Science
Gulbarga University
Gulbarga, Karnataka, India

Parashuram Bannigidad
Dept. of Computer Science
Gulbarga University
Gulbarga, Karnataka, India

Sai Geeta
Dept. of Computer Science
Gulbarga University
Gulbarga, Karnataka, India

ABSTRACT
The differential counting of white blood cell provides invaluable information to pathologist for diagnosis and treatment of many diseases manually counting of white blood cell is a tiresome, time-consuming and susceptible to error procedure due to the tedious nature of this process, an automatic system is preferable in this automatic process, segmentation and classification of white blood cell are the most important stages. The objective of the present study is to develop an automatic tool to identify and classify the white blood cells namely, lymphocytes, monocytes and neutrophil in digital microscopic images. We have proposed color based segmentation method and the geometric features extracted for each segment are used to identify and classify the different types of white blood cells. The experimental results are compared with the manual results obtained by the pathologist and demonstrate the efficacy of the proposed method.

Keywords
White blood cells, segmentation, image analysis, leukocytes, lymphocyte, monocyte, neutrophil, color segmentation.

1. INTRODUCTION
White blood cell composition reveals important diagnostic information about the patients. Substituting automatic detection of white blood cells for manually locating identifying and counting different classes of cells is an important topic in the domain of cancer diagnosis. Microscopic differential white blood cell count is still performed by haematologists, being indispensable in diagnostics with malignance suspicious. While its value as a reference method for blood samples containing abnormal cells remains indisputable, it is slow and subjective and its reproducibility is poor. Therefore, automation of this task is very helpful for improving the haematological procedure and accelerating diagnosis of many diseases [2].

There are three types of cells in normal human blood: red cells, leukocyte or white cells and blood platelets. Generally, red cells are simple and similar. While white cells contain nucleus and cytoplasm and there are different types of them. White cells are categorized into five groups: neutrophil, eosinophil, basophil, monocyte and lymphocyte. The texture, colour, size and morphology of nucleus and cytoplasm make differences among theses groups. In our paper we are considering only the nucleus. In blood smear, number of red cells is many more than white cells. For example, an image may contain up to 100 red cells and only 1 to 3 white cells. Platelets are small particles and are not clinically important. In laboratories, haematologists analyse human blood by microscope. Their main tasks in this area are: red cell count, white cell count and blood disorder detection. It is tedious task to locate, identify and count these classes of cells. Due to the importance of these processes, an automated system seems necessary and helpful. White cells are clinically more important than red cells and many of blood disorders are related to them. Thus, accurate segmentation of these cells is very important. Leukocytes count is used to determine the presence of an infection in the human body.

Lymphocytes are much more common in the lymphatic system. Lymphocytes are distinguished by having a deeply staining nucleus which may be eccentric in location, and a relatively small amount of cytoplasm.

Monocytes make up about 6% of white blood cells and have a somewhat unique and interesting role to play in human immune system. They have the kidney shaped nucleus and are typically agranulated. They also possess abundant cytoplasm. Monocytes are rather long lived compared to other white blood cells. They travel around in blood, looking for bacteria, viruses and other "waste" that needs removal. When they find something that needs cleaning up, they swallow the offending particle in a process known as "phagocytosis". After swallowing these bits, the monocyte will break the invader in to smaller pieces and present them on its cell surface so that passing T cells can "learn" more about the chemical make-up of the invader and make it easier to kill more of them.

Neutrophils defend against bacterial or fungal infection and other very small inflammatory processes that are usually first responders to microbial infection; their activity and death in large numbers forms pus. They are commonly referred to as polymorphonuclear (PMN) leukocytes, although technically PMN refers to all granulocytes. They have a multilobed nucleus which may appear like multiple nuclei, hence the name polymorphonuclear leukocyte. The cytoplasm may look transparent because of fine granules that are faintly pink. Neutrophils are very active in phagocytosing bacteria and are present in large amount in the pus of wounds. These cells are not able to renew their lysosomes used in digesting microbes and die after having phagocytosed a few pathogens [8].

The segmentation step is very crucial because the accuracy of the subsequent feature extraction and classification depends on the correct segmentation of white blood cells. It is also a difficult and challenging problem due to the complex nature of the cells and uncertainty in the microscopic image. Therefore, this step is the

most important challenge in many literatures and improvement of cell segmentation has been the most common effort in many research works.

Several researchers have previously proposed features to differentiate leukocyte (white blood cells) cells. The segmentation of leukocytes blood cells done by Mohommad Hamghalam and Ahmad Ayatollahi [1] based on histogram analysis and measurement of distance among nuclei. S. H. Rezatofighi et al.[2] have proposed the segmentation of white blood cell nucleus based on Gram-Schmidt orthogonalization process for amplifying the desired colour vectors. A variety of semi-automatic or automatic methods have been proposed to segment cell boundaries [3-6]. These methods include thresholding, watershed, nearest neighbourhood graphs, mean shift procedure and deformable models. The automatic classification of bacteria cells in digital microscopic images using simple shape geometric features is studied by Hiremath and Parashuram [7].

In this paper, we have proposed an automated segmentation, identification and classification of Leukocytes (White Blood Cells) namely, lymphocyte, monocyte and neutrophil in light microscopic images based on histogram equalization, thresholding and edge detection algorithms. The experimental results are compared with the manual results obtained by pathologist. The proposed method is more reliable and computationally less expensive.

2. MATERIALS AND METHODS

The microscopic images used in this study have been collected from the pathologist. A drop of blood is taken from the patient and smeared on a glass slide or cover slip. This specimen is Giemsa stained with clean 3 to 4 slides with water. Scrap the middle finger with 70% of alcohol & prick with the needle. Put a drop of blood on the slide. 1" inch from one end & spread with another slide. Note that the blood is spread over the slide if you don't get a even smear first. Repeat the procedure on clean slide. To get a smear of proper thickness, hold the spreading slide of an angle greater than 45^0. Draw on the two slides of the smear with a pencil cover the smear with the smearing by counting. The drops stain for 4min. then add same number of drops of distilled water to the stain. Let it stand for another 10 minutes. Blow gently on the air for every minute to make the solution mixture. Gently wash of stain under running water for 30 seconds and shake off the exceeds blow, dry with the slotting paper. The digital images of stained blood smear slide are captured by using a light microscopy imaging system at 40X magnification. We have considered 100 color images of leukocyte cells.

3. IMAGE ANALYSIS

Automatic recognition of white blood cells in light microscopic images usually consists of four major steps, including: pre-processing, image segmentation, feature extraction and classification. The pre-processing stage usually includes image enhancement of acquired image and is essentially performed in order to prepare the image for the vital segmentation stage. Individual objects of interest are separated from the background in the segmentation process. This is followed by a labelling operation (post-processing) in which, segmented objects of interest are tagged with unique labels that can be used to count the number of objects in the image. These labels along with spatial information of the segmented objects are used for the subsequent feature extraction procedure. The geometrical features are used to identify and classify the leukocyte cells, namely, lymphocyte, monocyte and neutrophil. The proposed method for the segmentation and classification of blood cell(leukocytes) is given below:

Algorithm 1: Training phase

Step 1 : Input the leukocyte colour cell image.
Step 2 : Convert the colour image into grayscale image.
Step 3 : Apply histogram equalization on grayscale image.
Step 4 : Perform pre-processing by using morphological operations, namely, erosion , reconstruction and dilation.
Step 5 : Segment the image of Step 4 by global thresholding and obtain resulting binary image.
Step 6 : Remove the border touching cells obtained in binary image and then perform labelling the segmented binary image.
Step 7 : For each labelled segment, compute geometric shape features (area, MajorAxislength/MinorAxislength, perimeter, circularity) and store them. Let a_i^k be the value of i^{th} parameter for k^{th} class. The i=1,2,3,4 correspond to area, MajorAxislength/ MinorAxis length, perimeter, circularity, respectively; and k=1,2,3 correspond to lymphocyte, monocyte, neutrophil, respectively.
Step 8 : Repeat Steps 1 to 7 for all the training images.
Step 9 : Compute minimum and maximum values of features of leukocyte cells, denoted by a^k_{imin} and a^k_{imax}, for all i and k, and store them as knowledge base.

Algorithm 2: Testing phase

Step 1 : Input the leukocyte colour cell image.
Step 2 : Convert the colour image into grayscale image.
Step 3 : Apply histogram equalization on grayscale image.
Step 4 : Perform pre-processing by using morphological operations, namely, erosion, reconstruction and dilation.
Step 5 : Segment the image of Step 3 using global thresholding and obtain resulting binary image.
Step 6 : Remove the border touching cells obtained in binary image and then perform labelling the segmented binary image.
Step 7 : For each labelled segment, compute geometric shape features a_i, i=1,2,3,4 .
Step 8 : Apply rule for classification of the leukocyte cells; if a_i lies in the range $[a^k_{imin}, a^k_{imax}]$, for i=1,2,3,4, then the cell (labelled segment) belongs to k^{th} class, where k=1,2,3 corresponds to lymphocyte, monocyte and neutrophil respectively.
Step 9 : Repeat the Steps 7 and 8 for all labelled segments and output the classification of identified leukocyte cells.

We have also proposed an improved algorithm for identification and classification of white blood cells in digital microscopic

images using color image segmentation method. The ratio of areas of nucleus and cytoplasm of a cell as a prominent feature is presented. For a given input image, the color image analysis is carried out based on HSV model. The experimental results are compared with the manual results obtained by pathologist. The performance of proposed algorithm is analyzed for four different feature sets [9], [10].

The input RGB image of leukocyte cell is converted into HSV color space and then only hue is considered. From the observation, it is clear that the hue value for the cell lies between 0.7 and 0.85. So, if the hue is between 0.7 to 0.85, then cell portion is extracted which contains the spurious regions along with it. These spurious regions can be eliminated by removing regions whose total number of pixels is less than T_A=1000. We obtain the actual cell region after removing the spurious regions. Now, nucleus has to be extracted from the already extracted cell region. Here, we consider only the saturation to extract the nucleus. Empirically, it is observed that, the nucleus has high saturation and it is above 0.45. Applying the above thresholding, it yields binarized images of nucleus and that of cell region. Finally, the cytoplasm region is obtained by subtracting binary image of nucleus from that of cell. For the experimentation, we use three feature sets, namely, F2, F3 and F4 for classification and compare with the results obtained for feature set F1 of the previous method:

F2=(area, eccentricity, equidiameter, perimeter, circularity)

F3=(area, eccentricity, equidiameter, perimeter, circularity,
ratio of areas of nucleus and cytoplasm)

F4=(area, majoraxislength/minoraxislength, perimeter,
circularity, ratio of areas of nucleus and cytoplasm)

The feature set F1=(area, majoraxislength/ minoraxislength, perimeter, circularity) is extended to F4 and F2 is extended to F3 by considering an extra feature, namely, the ratio of nucleus area and cytoplasm area.

The experimental results demonstrate that the proposed color image segmentation method using HSV model is efficient and effective in white blood cell segmentation, and provide better assistance to the pathologist in the process of differential counting and classification of white blood cells [12].

4. EXPERIMENTAL RESULTS

For the purpose of experimentation,100 light microscopic images of different types of leukocyte cells(non-overlapping) are considered which are taken from light microscopy. The implementation is done on a Intel core 2 duo processor @ 2.83 GHz machine using MATLAB 7. The input of leukocyte cell image is converted into grayscale image and then we perform histogram equalization and the morphological operations are applied. The resulting image is global thresholded to obtain segmented binary image. The segmented image is labelled and for each segmented region (known leukocyte cells), the geometric features are extracted. The Table 1 presents the geometric feature values computed for the segmented leukocyte cells, namely, lymphocytes, monocytes and neutrophil of the image in Figure 1(c), (f) and (i). The minimum and maximum values of the geometric features (F2, F3) and (F1, F4) of these cells in the training images are stored in the knowledge base of the leukocyte cells and are given in the Table 2 and Table 3, respectively. Some sample training images of leukocyte cells are shown in Figure 2.

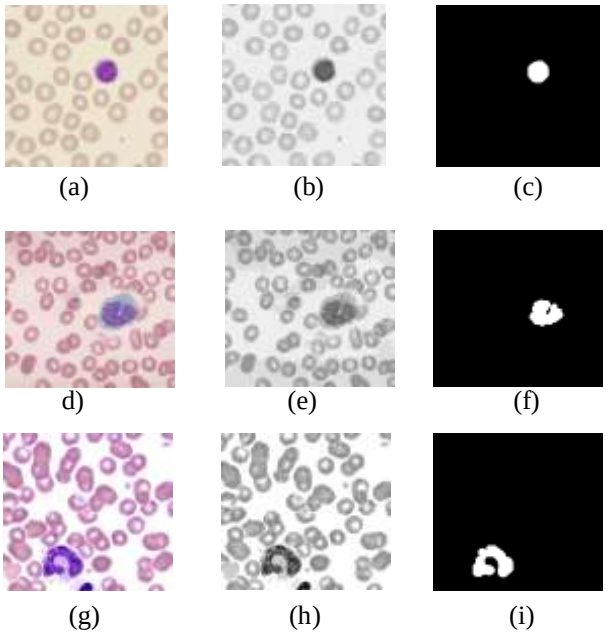

Figure. 1: (a)Giemsa stained colour image of lymphocyte (b) Gray scale image of lymphocyte (c) Segmented image of lymphocyte (d)Giemsa stained colour image of Monocyte (e) Gray scale image of monocyte (f) Segmented image of monocyte (g) Giemsa stained colour image of Neutrophil (h) Gray scale image of neutrophil (i) Segmented image of neutrophil.

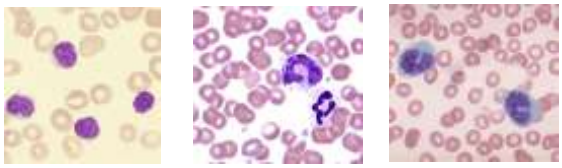

Figure 2: Sample training images of leukocyte cell images.

In the testing phase, the feature extraction is applied for each segment and the test feature values are compared with the knowledge base for each type of leukocyte cells, namely, lymphocytes, monocytes and neutrophil, for classification using minimum-maximum range. The proposed method is computationally less expensive. The classification performance of the proposed method by using different feature sets F1, F2, F3 and F4 of leukocyte cells is given in the Table 4. It yields classification rate in the range 92% to 98% for different leukocyte cells based on feature set F1, and still better rates in the range 98% to 99% in case of feature set F4. The feature set F4 yields better performance, despite containing only five geomteric features, as compared to other feature sets. The Figure 3 shows some sample test images used for classification of different leukocyte cells, namely, lymphocyte, monocyte and neutrophil.

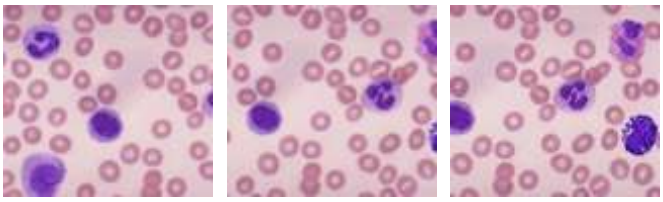

Figure 3: Sample test images of leukocyte cell images

Table 1. The geometric feature values (F1) of the cell, regions of the images in Figure 1(c), (f) and (i)

Cell Types	Area	MajorAxisLength/ MinorAxisLength	Perimeter	Circularity
Lymphocytes	883	1.0605	128	0.67734
Monocyte	1340	1.2850	220	0.34796
Neutrophil	1806	1.3428	314	0.23021

Table 2. The geometric feature values (F2, F3) of the cell, regions of the images in Figure 1(c),(f) and (i)

Leukocyte Cell Type		Area	Eccentricity	EquiDiameter	Perimeter	Circularity	Ratio of areas of nucleus and cytoplasm
Lymphocyte	Min	724	0.321	30.35	118	0.500	1.1997
	Max	1059	0.604	36.72	155	0.666	1.690
Monocyte	Min	722	0.524	30.32	167	0.225	0.8676
	Max	1756	0.826	47.28	254	0.499	1.2208
Neutrophil	Min	716	0.51	30.01	162	0.249	0.41734
	Max	954	0.66	35.01	206	0.456	1.481

Table 3. Maximum and minimum values of features (F1, F4) of leukocyte cells, namely, Lymphocyte, monocyte and neutrophil

Leukocyte cell types		Area	MajorAxisLength/ MinorAxisLength	Perimeter	Circularity	Ratio of areas of nucleus and cytoplasm
Lymphocyte	Min	869	1.0353	128	0.59250	0.8475
	Max	1277	1.4625	162	0.67964	1.4125
Monocyte	Min	1020	1.1187	172	0.33370	0.4165
	Max	2039	1.5655	252	0.45329	1.3711
Neutrophil	Min	816	1.1583	176	0.23021	0.8484
	Max	1821	1.6330	314	0.44063	1.4175

Table 4. Classification rates for the different leukocyte cells using feature sets F1, F2, F3 and F4

Leukocyte cell images	No. of cells in test images	Classification rate in (%)			
		F1	F2	F3	F4
Lymphocyte	34	98%	98%	98%	98%
Monocyte	12	92%	96%	99%	99%
Neutrophil	29	95%	97%	99%	99%

5. CONCLUSION

In this paper, we have proposed an automated image segmentation and classification of electron microscope images and extracting geometric features of leukocyte cells. The experimental results are compared with the manual results obtained by pathologist. The proposed method is more reliable and computationally less expensive and yet yields comparable classification rate in the range 92% to 98% for different leukocyte cells based on feature set F1, and still better rates in the range 98% to 99% in case of feature set F4. It could be improved further by better pre-processing methods and feature sets, which will be taken up in our future work. This method of identification and classification of leukocytes can also be extended to two other cell types, namely, eosinophil and basophil.

6. ACKNOWLEDGEMENTS

The authors are indebted to the referees for their valuable comments and suggestions. Further, the authors are also grateful to Dr. Anand Nagalikar, Pathologist, KBN Medical College & Hospital Gulbarga, for providing Leucocyte cell images and manual results of the cell images by visual inspection.

7. REFERENCES

[1]. Mohammad Hamghalam and Ahmad Ayatollahi, "Automatic Counting of Leukocytes in Giemsa-Stained Images of Peripheral Blood smear" IEEE International Conference on Digital Image Processing, pp. 13-16, 2009.

[2]. S. H. Rezatofighi, H. Soltanian-Zadeh, R. Sharifian and R. A. Zoroofi, "A New Approach to White Blood Cell Nucleus Segmentation Based on Gram-Schmidt Orthogonalization" IEEE International Conference on Digital Image Processing, pp. 107-111, 2009.

[3]. Saif Zahir, Rejaul Chowdhury, and Geoffrey W. Payne, "Automated Assessment of Erythrocyte Disorders Using Artifical Neural Network", IEEE International Symposium on Signal Processing and Information Technology, 2006.

[4]. Silvia Halim, Timo R. Bretschneider, Yikun Li, Peter R. Preiser and Claudia Kuss, "Estimating Malaria Parasitaemia from Blood Smear Images, IEEE ICARCV 2006.

[5]. Refai, H., Li, L., Teague, T. K., and Naukam, R., "Automatic Count of hepatocytes in microscopic images" Proc. Int'l Conf. on Image Processing, 2, pp. 1101-1104, Sept 2003.

[6]. FANG Yi, ZHENG Chongxun, PAN Chen and LIU Li, "White Blood Cell Image Segmentation Using On-line Trained Neural Network", Proceedings of the IEEE Engineering in Medicine and Biology 27th Annual Conference Shanghai, China, September 1-4, 2005.

[7]. P.S. Hiremath and Parashuram Bannigidad , Automatic Classification of Bacterial Cells in Digital Microscopic Images, ICDIP 2010, February 26-28, 2010 Singapore (Accepted).

[8]. http://en.wikipedia.org/wiki/White_blood_cell

[9]. F. Kurugollu, B. Sankur. A. E. Harmanci, "Color Image Segmentation using histogram multithresholding and fusion," Image and Vision computing 19, pp. 915-928, 2001.

[10]. Sawsan F. Bikhet, Ahmed M. Darwish Hany A. Tolba, and Samir I. Shaheen, "Segmentation and Classification of White Blood Cells " IEEE pp. 2259-2261, 2000.

[11]. Vincenzo Piuri, Fabio Scotti. Morphological Classification of Blood Leukocytes by Microscope images, 2004. IEEE International conference on computational Intelligence for Measurement Systems and Applications Boston. MD, USA, 14-16 July 2004.

[12]. Rastislav Lukac, Konstantinos, N. Plataniotis, Color Image Processing, Methods and Applications, CRC Press(2007)