International Journal of Research in Medical Sciences
Sanyal P et al. Int J Res Med Sci. 2019 Mar;7(3):871-875
www.msjonline.org
pISSN 2320-6071 | eISSN 2320-6012

DOI: http://dx.doi.org/10.18203/2320-6012.ijrms20190939

Original Research Article

The watershed transform in pathological image analysis: application in rectiulocyte count from supravital stained smears

Parikshit Sanyal*, Sanghita Barui

Department of Pathology, MH Jalandhar, Punjab, India

Received: 02 January 2019
Accepted: 30 January 2019

***Correspondence:**
Dr. Parikshit Sanyal,
E-mail: cmacus@gmail.com

ABSTRACT

Background: Morphometric studies based on image analysis are a useful adjunct for quantitative analysis of microscopic images. However, effective separation of overlapping objects if often the bottleneck in image analysis techniques. We employ the watershed transform for counting reticulocytes from images of supravitally stained smears.
Methods: The algorithm was developed with the Python programming platform, using the Numpy, Scipy and OpenCV libraries. The initial development and testing of the software were carried out with images from the American Society of Hematology Image Library. Then a pilot study with 30 samples was then taken up. The samples were incubated with supravital stain immediately after collection, and smears prepared. The smears were microphotographed at 100X objective, with no more than 150 RBCs per field. Reticulocyte count was carried out manually as well as by image analysis.
Results: 600 out of 663 reticulocytes (90.49%) were correctly identified, with a specificity of 98%. The major difficulty faced was the slight bluish tinge seen in polychromatic RBCs, which were inconsistently detected by the software.
Conclusions: The watershed transform can be used successfully to separate overlapping objects usually encountered in pathological smears. The algorithm has the potential to develop into a generalized cell classifier for cytopathology and hematology.

Keywords: Automated, Cell classification, Image analysis, Reticulocyte, Watershed transform

INTRODUCTION

Image analysis is one of the modalities that impart a strong and objective foundation to the study of pathology. In recent times, in addition to immunohistochemistry, flow cytometry and molecular genetic studies, image analysis has emerged at the forefront of new developments in diagnostic techniques. However, unlike the aforementioned techniques, the data analysed in image anlaysis is not discrete, but continuous and overlapping in nature. Thus, the field of image analysis in pathology has taken a very long time to mature. In the present study, we have attempted to develop an algorithm to count reticulocytes from microphotographs of a supravitally stained peripheral blood smear. The algorithm has the potential to develop into a cell classifier for all blood cell types.

Classification of cells by image analysis has been extensively studied. Back in 1979, Bacus et al developed and patented a computerized cell classifier for cytological images.[1] Classification systems for cervical smears, both

liquid based and conventional, by Image Analysis modalities have been developed.[2,3] However, in recent times, research on image analytic techniques in Hematology has also gained momentum. An image classifier has also been developed by Patale et al for acute myeloid leukemia and by Nikitaev et al.[4,5] A similar tool has also been developed for acute lymphoblastic leukemia.[6]

A method of reticulocyte count by image analysis was evaluated by Hackney et al, who concluded that image analysis is a viable tool for reticulocyte analysis in high volume laboratories.[7] In a study later in 1992, Schimenti et al demonstrated that image analysis has better precision than manual counting but less than flow cytometry.[8] Over time, with improvements in automation, flow cytometry has now superceded manual methods as being the gold standard for reticulocyte counts. However, newer technologies such as machine learning algorithms, provide a robust backbone to the Automated Image Analysis procedure. The present study is the initial step in the development of a generalized image analysis based cell classifier. Reticulocytes being easily discernible, is also the fastest for machines to learn and identify.

The aim of this study was to develop and test a watershed transform based algorithm for counting reticulocytes from microphotographs of supravitally stained peripheral blood smears.

METHODS

Initial development of software framework

A set of standard images were chosen from ASH Image Library and standard text books (E-format). for purpose of development and standardization of the software. The software was developed using the Python programming language, the Numpy mathematical and statistical library, Open CV and Scikit-image analysis library. Image analysis techniques like thresholding, Gaussian blur and Watershed segmentation was used for segmentation of red blood cells and identification of reticulin granules. The python source code is attached as Appendix.

The initial step is to segment the image in several connected regions which share at least one common pixel. The software then separates red blood cells through successive filters (Figure 1). The first transformation is a threshold, which has been set depending on the background illumination of the image. The software then applies a Watershed transform to separate red blood cells that are touching or overlapping each other. A certain size threshold and nuclear complexity threshold is applied to exclude white blood cells. The software then calculates number of reticulin dots depending on the intensities of the red, blue and green channels in the red blood cells.

The stain used for reticulocytes was a mix of new methylene blue (20mMol/L), citric acid (180mMol/L) and sodium chloride (1.55Mol/L) in phosphate buffer (0.1mMol/L), manufactured by BIOLAB diagnostics. 0.2ml of freshly collected blood samples in EDTA was mixed with 0.2ml of reticulocyte stain and incubated for 20minutes at room temperature. Smears were made, air dried and examined under oil immersion.

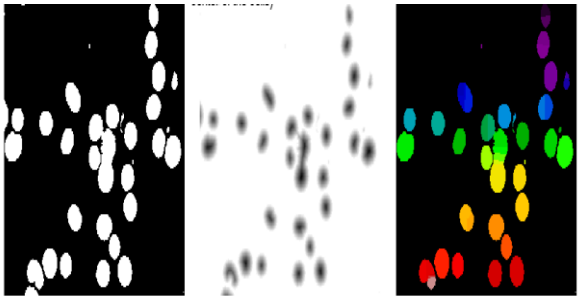

Figure 1: The software transforms the image into greyscale, performs a distance transform and watershed to separate red blood cells. A) Thresholded black and white, B) Distance transform (pixel intensities are inversely proportional to distance from the center of cell, C) watershed transform for cell separation.

Deployment

The parameters for image segmentation was set to classify microscopic foci from a set of reticulocyte smear images, collected from archives of this hospital photographed from the stained smears, after reporting out and archiving. All photographs were taken within 24hours of preparation of smear. Random foci containing 100-150 red blood cells in one oil immersion field were photographed from each smear. Manual reticulocyte count was performed in the same fields. The microscope used for this purpose was Labomed ATC 2000.

All of the images were photographed at 100x magnification, after correction of white balance, at the same condition of illumination. After digital archiving, the dimension of the images were resized and cropped to less than or equal to 640x480 pixels in size carefully preserving the aspect ratio.

Using standardization data from the development phase, the parameters for analysis like maximum cell size and illumination threshold. The software was then run on the set of images collected from the Hospital archives. The criterion used for identification of a reticulin dot was as follows. If the blue channel was sufficiently dark (<210) AND the staining was intense (red channel <70), then difference between blue and green channel must be >20. For lighter staining (red channel >70), the difference should have been >20. An RBC must have >3% of area to be reticulin to be classified as a reticulocyte. These parameters were calibrated for the particular microscope in use and its illumination setting.

RESULTS

Table 1 shows the results of analysis of the photomicrographs. 3305 red blood cells were detected by the software. Manual reticulocyte count on these images showed 663 reticulocytes in total, while the software could detect 643 of them. Thus, the detection rate of reticulocytes was 643/663 = 90.49%.

Table 1: Metrics from 30 photomicrographs in the study.

	RBC automated count	Retic automated count	Retic manual count	Difference	False positive	False negative
1	107	3	5	2	0	2
2	114	30	26	-4	4	0
3	138	27	23	-4	4	0
4	116	23	23	0	0	0
5	99	15	20	5	0	5
6	94	24	30	6	0	6
7	135	25	38	13	0	13
8	99	11	20	9	0	9
9	117	28	30	2	0	2
10	134	36	32	-4	4	0
11	128	32	28	-4	4	0
12	127	10	9	-1	1	0
13	128	34	40	6	0	6
14	112	32	36	4	0	4
15	127	38	42	4	0	4
16	136	34	37	3	0	3
17	138	53	48	-5	5	0
18	121	30	32	2	0	2
19	120	29	30	1	0	1
20	112	29	27	-2	2	0
21	28	9	7	-2	2	0
22	34	6	8	2	0	2
23	128	14	12	-2	2	0
24	106	13	8	-5	5	0
25	122	15	10	-5	5	0
26	96	9	7	-2	2	0
27	88	7	7	0	0	0
28	92	3	6	3	0	3
29	98	10	7	-3	3	0
30	111	14	15	1	0	1
Total	3305	643	663	20	43	63

Total error in detecting reticuloyctes over 30 microphotographs was 20, i.e. the summation of reticulocytes that were missed in each field as well as those RBCs falsely detected as reticulocytes. Correlation between manual and automated reticulocyte count was only good (r = 0.90, regression equation y = 0.981x + 1.055) (Figure 2). Paired differences between manual and automated reticulocyte counts showed a mean error of 0.68 reticulocytes per field (Figure 3).

A number of false positive results (43) sprang up during the analysis, most of them being RBCs with a bluish tinge but without definitive reticulin granules. Color images are represented by computers as an array of three values, one for each of the colors red, green and blue. The software operates based on detection of individual reticulin granules, by virtue of the difference of the pixel values in the green and blue channel. The software also takes into account the amount of reticulin granules in the RBC. It is a matter of choosing the right cutoff values for two parameters so that detection rate would increase.

- The difference in value of the blue channel and the green channel in a reticulin granule,
- The percentage of reticulin areas in an RBC to make it qualify as a reticulocyte.

However, any increase in detection would also result in increased false positives. By method of trial and error, we found that a value of 40 for parameter (a) and 3% for parameter (b) above produces the maximum detection rate.

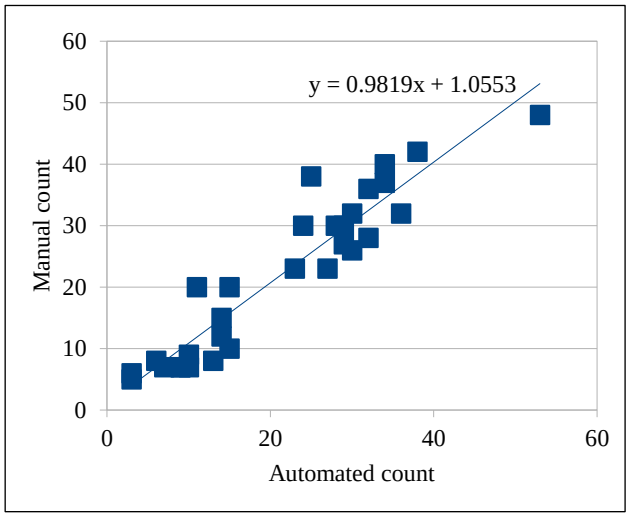

Figure 2: Plot of manual vs automated reticulocyte count.

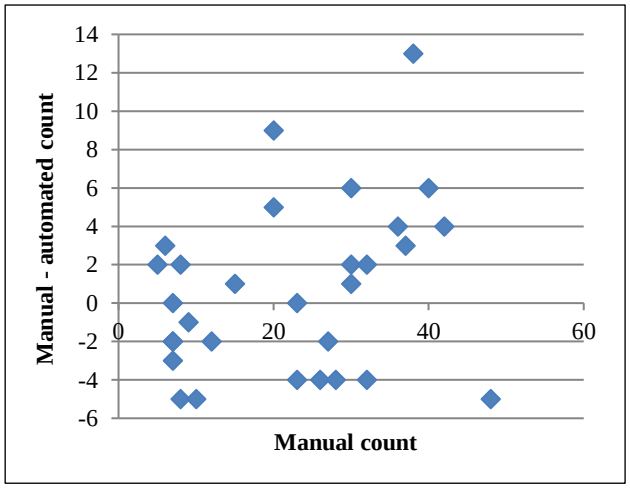

Figure 3: Paired difference between manual and automated count.

DISCUSSION

The standard technique for reticulocyte count is sating with fluorescent dyes, a technique which has been later incorporated into flow cytometry.[9-11] The present study aims not to replace flow cytometry as the gold standard tool, but to evaluate an image analysis based software as a potential automated cell classifier.

The sensitivity of detection of reticulocytes in the present study was 90.49%. Among a total of 663 reticulocytes, 600 were correctly identified by the software. The reticulocytes that were missed were due to improper segmentation of closely touching RBCs, such as in Figure 4. Correlation coefficient achieved by the present software is 0.90.

Hackney et al have reported a higher correlation coefficient (r = 0.96), after following a similar method of manual and automated count using the Hematrak 590.5 Their method of staining included incubation with equal volumes of new methylene blue over 10 to 15 minutes. The manual method was performed by counting the number of reticulocytes during a count of 1,000 red blood cells. Criteria for identification of reticulocytes were defined by the NCCLS Proposed Standard H16P: two or more clumps or discrete granules of "reticulin". The Hematrak counted 1000 RBCs per slide. They found the manual count to be imprecise and inaccurate (CV 44.2%), and the automated method was much more precise (CV 25%). They also reported that there is a slight positive bias of the automated methods at low reticulocyte counts, a result which has been reproduced in our study. This effectively translates to higher error rate when the field contains very few reticulocytes.

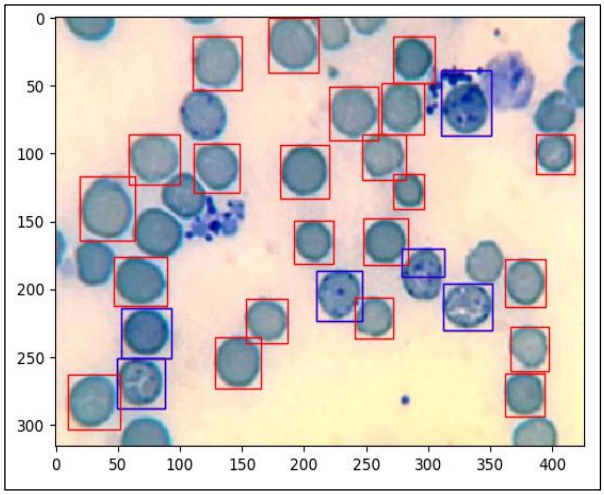

Figure 4: RBEs detected by the software marked in red, reticulocytes in blue; the WBC has been excluded by the software.

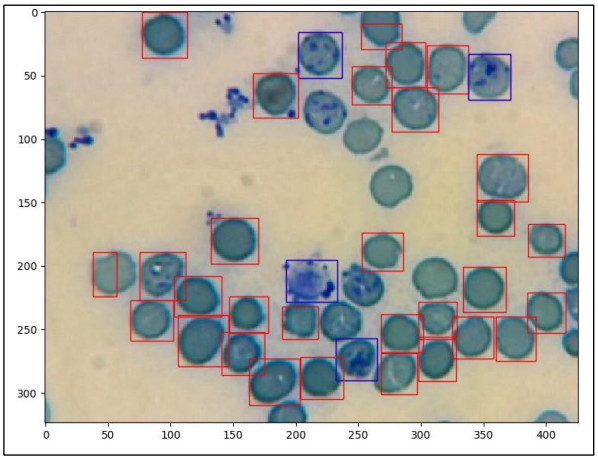

Figure 5: WBC has been detected as a reticulocytes.

Schimenti et al performed reticulocyte counting by flow cytometry with thiazole orange as well as manual or automated counting of new methylene blue stained blood smears, in a method similar to ours, but with lesser incubation time (10 minutes).[8] 338 Smears were prepared using an automatic blood-smearing device (Miniprep). Five hundred RBC were counted by two different technologists. They consistently demonstrated that thiazole orange is more sensitive than new methylene blue because the window of analysis includes an increased fraction of mature reticulocytes. Precision, from high to low, was flow cytometry > image analysis >manual counting. Unlike the study by Schimenti et al, no comparison with flow cytometry was carried out, and manual count was used as the Gold Standard procedure.[8]

Both of the two aforementioned studies have compared cumulative reticulocyte counts per 500-1000 red blood cells, counted over multiple high power fields. In a significant departure from their method, we have compared manual and automated methods on a 'reticulocytes per high power field basis', and not as an overall count over many fields. Thus, the same set of slides were used for both manual and automated counting, unlike Schimenti et al. We have used a field by field comparison of the two methods on the same slide, to better track the performance of the software, and also get a continuous visual feedback of the process at runtime (Figure 1).

White blood cells (WBCs) were excluded from the count both by a size and a circularity criteria, as in Figure 4. However, an occasional WBC nucleus was picked up as false positive (Figure 5). Majority of false positives were RBCs with a diffuse bluish tinge but without reticulin granules, such as Figure 4. Excluding these false positives will require further refinement of algorithm.

The criteria chosen for identification of reticulocytes was chosen arbitrarily, to achieve a satisfactory level of sensitivity and specificity. The criteria can be adjusted to increase sensitivity or specificity, but not both at the same time. The criteria have been calibrated to suit one particular microscope and illumination conditions. For a different microscope and camera combination, the criteria need to be re adjusted.

Reticulocytes were chosen for the initial development of the software for their easy recognizability and relatively low cellular complexity. The software may be having the potential to evolve into a generalized cell classifier for cytopathology and histopathology.

Funding: No funding sources
Conflict of interest: None declared
Ethical approval: Not required

REFERENCES

1. Bacus JW, inventor; Rush-Presbyterian-St Luke's Medical Center, assignee. Dual resolution method and apparatus for use in automated classification of pap smear and other samples. United States patent US 4,175,860. 1979 Nov 27.
2. Sanyal P, Ganguli P, Barui S, Deb P. Pilot study of an open-source image analysis software for automated screening of conventional cervical smears. J Cytol. 2018;35:71-4.
3. Plissiti MR, Charchanti A, Krikoni O, and Fotiadis DI. Automated segmentation of cell nuclei in pap smear images. In ITAB 2006-Proceedings of Informational Technology in Biomedicine; 2006.
4. Patale SS, Pansambal BH. Acute myeloid leukemia detection in blood, microscopic image by using PNN. Int J Innovative Res Computer Communication Eng. 2016;4(4):7094.
5. Nikitaev VG, Nagornov OV, Pronichev AN, Polyakov EV, Dmitrieva VV. The blood smear image processing for the acute leukemia diagnostics. Int J Biol Biomed Engg. 2016;10:109.
6. Amin MM, Kermani S, Talebi A, Oghli MG. Recognition of acute lymphoblastic leukemia cells in microscopic images using k-means clustering and support vector machine classifier. J Medical Signals Sensors. 2015 Jan;5(1):49.
7. James R. Hackney JR, Cembrowski GS, Prystowsky MB, Kant JA. Automated reticulocyte counting by image analysis and flow cytometry. Lab Med. 1989:551.
8. Schimenti KJ, Lacerna K, Wamble A, Maston L, Iaffaldano C, Straight M, et al. Reticulocyte quantification by flow cytometry, image analysis, and manual counting. Cytometry. 1992;13(8):853-62.
9. Tanke HJ, Rothbarth PH, Vossen JM, Koper GJ, Ploem JS. Flow cytometry of reticulocytes applied to clinical hematology. Blood. 1983 Jun;61(6):1091-7.
10. Seligman PA, Allen RH, Kirchanski SJ, Natale PJ. Automated analysis of reticulocytes using fluorescent staining with both acridine orange and an immunofluorescence technique. Am J Hematol. 1983 Feb;14(1):57-66.
11. Metzger DK, Charache S. Flow cytometric reticulocyte counting with thioflavin T in a clinical hematology laboratory. Arch Pathol Lab Med. 1987;111(6):540-4.

Cite this article as: Sanyal P, Barui S. The watershed transform in pathological image analysis: application in rectiulocyte count from supravital stained smears. Int J Res Med Sci 2019;7:871-5.