ACCURATE RED BLOOD CELLS AUTOMATIC COUNTING IN MICROSCOPIC THIN BLOOD SMEAR DIGITAL IMAGES

Naveed Abbas,. Dzulkifli Muhamad

ViCube Research Lab, Faculty of Computing, University Technology Malaysia, Skudai, Johar Bahru, Malaysia. naveed23a@yahoo.com, dzulkifli@utm.my

Abstract: The aim of this study is to automate the counting process of Red Blood Cells in thin Blood smear images in more accurate, efficient and universal way. The Red Blood Cells have important role in the blood; their counting is part of the complete blood count test and is frequently suggested by the Physician because the Red Blood Cells have hemoglobin, responsible to carry oxygen to various tissues of the body. The number of Red Blood Cells both (Low and High) deviations from normal range is an important indicator about any disorder existence in the body. At present mostly the counting process is performed manually which is laborious, error prone and time consuming. The automated diagnosing gain the attention of the researchers from the last two decades because it assist the experts to reduce the burden of errors, labor and time of examination. In this regard, too much research has been performed on the automation of the counting process of the Red Blood Cells but still the test demands to be done in a proper, efficient, accurate and realistic way. The proposed method achieved an average True Positive Rate (TPR) of 94%, True Negative Rate (TNR) of 6%, average accuracy of 97% and average error rate of 3%.

Erythrocytes counting, Automatic Diagnosing, Rouleaux splitting.

1. INTRODUCTION

The Red Blood Cells count is demanding in various blood tests because the deviation of number of Red Blood Cells from normal range in both cases (Low and High) is an important indicator about any disorder in the body. The normal range of the Red Blood Cells in male is 4.7-6.1million cells /mcl and in female is 4.2- 5.4 million cells/mcl. The high number of RBCs then normal range indicates Kidney tumor. Heart diseases. Low Blood oxygen level etc. while the low number of Red Blood Cells from its normal range indicates, Anemia, Hemorrhage, Leukemia, Malnutrition, Nutritional deficiencies like iron, folate, copper etc.[1]. Due to consumption of too much time, jeopardy of errors and much physical and mental labor on the part of hematologist increases the demand of automatic counting techniques to combat the mentioned problems by assisting the hematologists. In this connection, many researchers did much work but still the work needs to be more efficient, robust, accurate and realistic. This study considered the proposed technique in the context that it will be efficient, accurate, robust and realistic. Counting Red Blood Cells through image processing techniques is not difficult task but for high accuracy it involves several other problems i.e. image pre-processing, separation of single and clustered Red Blood Cells. If theses mentioned problems are not addressed in proper ways then the accuracy will be compromised because the rouleaux or clustered Red Blood Cells are appeared as a single area and in reality it is combination of more than one Red Blood Cell. This study considered all these problems and after solving the given problems then count the Red Blood Cells.

Recently, too much effort have been made by researchers to develop algorithms for counting of Red Blood Cells addressing the problems of splitting the clustered Red Blood Cells and show a high degree of success but still needs improvements to address the mentioned hurdles in proper way. The study made by [2], the authors mentioned that counting Red Blood Cells is not a big issue in image processing but the hurdles like clustered Red Blood Cells splitting is too important because they will affect the accuracy that's why they did it through concavity points finding and splitting. However, they did not mention how to separate the single and clustered Red Blood Cells and identification of clusters existence while the Red Blood Cells are counted using boundaries tracing and labeling. In the study, of [3], the authors did not consider the separation and clustered Red Blood Cells splitting but did the counting. Red Blood Cells counting without solving the problem of cluster Red Blood Cells Splitting compromise on the accuracy. Some studies while counting the Red Blood Cells do not consider the clumps and overlaps of Red Blood Cells for splitting but they rely on guessing Area based estimation approaches as mentioned in the work of [4] and [5]. The problem in this approach is that in some cases we want to note the disorder as well in the Red Blood Cell in such case this approach will fails while also the areas of Red Blood Cells by most of the studies considered as circular, which is not true as because morphology of the Red Blood Cells highly changes due to any disorder. Circular Hough Transform based approaches for counting and splitting as mentioned by [6,7,8,9] mainly considered the Red Blood Cells as circles which is not true because Red Blood Cells morphology is not static and changed by other diseases. The approaches adopted by previous studies to combat the problem of clumped and overlapped Red Blood Cells splitting are divided into the following categories i.e. Morphological operation based includes erosion, dilation or opening closing to split the clusters of Red Blood Cells [10,11,12]. However, the main problem in morphological based approach is that it works well in overlap of Red Blood Cells not more than two cells but in reality we have some clumps which are very long chains. Concavity based approaches deal the problems in the way to find out the concavity regions and some cases the concavity points and split the clustered Red Blood Cells through lines cuts or circles drawing or ellipses drawing as stated in the studies of

[13,14,15,16,17,18,19,20,21]. The concavity based approaches gives good results but in some cases they are computationally very expensive. Watershed based techniques includes all form of watershed algorithm based etc as presented by the studies of [22,23,24,23,25,26,27]. Watershed based approach have certain degree of success but in dense clumps it results in over segmentation while in some cases also suffered from the problem of under segmentation. Edges or contour based techniques can gives solution in the form of analyzing split edges and linkages of contours etc as mentioned in the works of [28,29,30]. This approach working well but required model based on some templates and complex both in execution as well as in implementation. Model based approach gives various models in the form of circles through various theories like Gestalt, geometrical theories etc as presented in the work of [31,32,33]. The problem in this approach seems to be unrealistic as due to its highly complex nature and implementation. Also it is too much expensive computationally.

2. MATERIALS AND METHODS

In this paper we performed the experimentations on Microscopic thin blood smear digital images set of 20 images, which are obtained from the [34], which are free available for research purposes. The proposed methodology started with image pre-processing then the slide image is then passed from separation of single and clustered Red Blood Cells for the purpose to improve the efficiency, then the clustered Red Blood Cells is passed from splitting process because without splitting the accuracy is compromised. This whole process is presented as overall methodology of this study in Figure 1and its simulated diagram in the form of images is depicted in Figure 2.

2.1 Image Pre-processing

As image pre-processing we only convert the input RGB image to binary image through Global thresholding OTSU for the purpose to reduce the processing time.[35,36] After conversion small areas are identified as noise and removed from the binary image and holes in the centers of the Red Blood Cells, formed due to hemoglobin in the centers of the Red Blood Cells and its similarity to the background are filled and we get the image presented in Figure 3 which is ready for further processing.

2.2 Separation of Single and Clustered Red Blood Cells

In separation of clustered Red Blood Cells and single Red Blood Cells we applied a double check on the convex hulls (through equation 1) of all the Red Blood Cells. We find the areas and elongation of the convex hulls of the Red Blood Cells as mentioned in equation 2 and 3 respectively. Next between these two measures we consider median to find median area and elongation as mentioned in equation 4.

we consider median among many central tendency measure for the purpose that the median is the best central tendency measure in case when the data values are irregular and having some small while some large values. We divide the area of every convex hull of Red Blood Cells with the median area, the result obtained if equals to 1 or near to 1 are considered as single Red Blood



Figure 1 Overall Methodology



Figure 2 Simulated Images of the whole process



Figure 3 Matlab Results a) Original RGB Image b) Pre-processed Image

cells and are considered for mask of single Red Blood Cell while the negation of the single Red Blood Cell resulted in multi-Red Blood Cells mask. Then we pass the single Red

July-August, 2014

Blood Cell mask to the pixel IDX list of the input image and obtained the image of single Red Blood Cells while on passing the multi-mask we obtained the image of clustered Red Blood Cells. In the same way we performed for the second check but instead of area we used elongation here. $\sum_{i=1}^{|X|} \alpha_i x_i |\langle \forall_i : \alpha_i \ge 0 \rangle \& \sum_{i=1}^{|X|} \alpha_i = 1$ (1)

where , $|X| = \text{finite set of points}, x_i \text{ is point } |X|$ while α_i is weight assigned to x_i , the sum of the weights must be equal to 1 mean normalized.

Area = No. of Pixels

No.of Pixels= Pixels defining the convex hull object of Red Blood Cells.

(2)

Elongation = $\frac{\text{Length}}{\text{Breadth}}$ (3) where, Length = Major Axis and Breadth = Minor Axis $\sigma^2 = \frac{(X-\mu)^2}{N}$ (4)

where, X represents the area in one case while elongation in the other case, N is the number of terms in distribution.

2.3 Splitting Clustered Red Blood Cells

After separation of single and clustered Red Blood Cells, the image of clustered Red Blood Cells is further considered for splitting the clusters into single cleaved Red Blood Cells. In the splitting process we first trace the boundaries of all clustered Red Blood Cells, first we divide the boundary into two halves using equation 5, then taking the first point of the boundary as P1(x1,y1) while, P2(x2,y2) is the last point of the first half of the boundary. After, finding the points P1 and P2 we calculate the distance between P1 and P2 using equation 6, once find out the distance the next process is to divide the boundary according to the number of Red Blood Cells in each cluster and take these division points and the distances between the consecutive points using equation 6. The same points are marked on the other half of boundary and then using Digital Differential Analyzer graphics algorithm to draw lines in between respective end points and in this way after applying a slight erosion(Morphological operator) the occlusions are cleaved into single Red Blood Cells. The idea is simulated in the diagram depicted as Figure 3.

$$Index = \frac{Length(Boundary)}{2}$$
(5)

where, boundary is the boundary of clumped or overlapped RBCs and index is the index of boundary containing its points.

$$D = \sqrt{\frac{(x_2 - x_1)^2 + (y_2 - y_1)^2}{2}}$$
(6)

$$No.of.Parts = \frac{D}{No.ofRBCs}$$
(7)

where, Number of RBCs we can found while dividing the convex hull area by the median area of single RBC.



1121

Figure 4 The simulated diagram of the concept of splitting RBCs

2.4 Counting of Red Blood Cells

Once the Clustered Red Blood Cells are cleaved into single Red Blood Cells then it is not difficult to count them. Thus for counting we consider the Matlab built-in function bwlabel, which uses a binary image and produces a label matrix L having value 0 for the background pixels while gives greater integer values than 0 according to the number of objects in a fashion that assign 1 to the first object, assign 2 to the second object and in this way increase the number according to the number of objects in an arbitrary order.



Figure 5 Matlab Results a) Presents Original RGB Image b) Presents the single RBCs separated from Clustered RBCs which are presented in c) while d) Presents their counting with number labels on each cleaved RBC

3. RESULTS AND DISCUSSION

In this section we are presenting the results obtained from the implementation of the above concepts through both qualitatively and quantitatively on microscopic thin blood smear digital image dataset of 20 images obtained from DPDx [34].

3.1 Qualitative Analysis

In Qualitative analysis we performed the experimentation on the images having clustered Red Blood Cells and we successfully cleaved the clustered Red Blood Cells into single Red Blood Cells and then performed the counting which will increase the accuracy as presented visually with ground reality.



Figure 6 Matlab Results of Slide 1 a) Presents Original RGB Image b) Presents Single RBCs separated from clustered RBCs along with the number labels for counting c) Presents cleaved RBCs with

number labels on each RBC



Figure 7 Matlab Results of Slide 2 a) Presents Original RGB Image b) Presents Single RBCs separated from clustered RBCs along with the number labels for counting c) Presents cleaved RBCs with number labels on each RBC



Figure 8 Matlab Results of Slide 3 a) Presents Original RGB Image b) Presents Single RBCs separated from clustered RBCs along with the number labels for counting c) Presents cleaved RBCs with number labels on each RBC



Figure 9 Matlab Results of Slide 4 a) Presents Original RGB Image b) Presents Single RBCs separated from clustered RBCs along with the number labels for counting c) Presents cleaved RBCs with number labels on each RBC

3.2 Quantitative Analysis

In this category we performed the experimentation on the images data set of 20 images and compared the counted Red Blood Cells results with the manually counted Red Blood Cells results. The Red Blood Cells are manually counted by experts. Using the confusion matrix we calculate the sensitivity or True Positive Rate (TPR) or Recall, Accuracy (AC), Error Rate (Er.R) and Specificity or True Negative Rate (TNR) with equations mentioned as equations 8, 9, 10 and 11 respectively.

$$TPR = \frac{A}{A+B} \tag{8}$$

$$AC = \frac{A+D}{A+B+C+D} \tag{9}$$

$$Er.R = 1 - AC \tag{10}$$

$$TNR = \frac{D}{C+D} \tag{11}$$

Confusion Matrix		Detected		
		Positive	Negative	
Actual	Positive	A: True +ve	B: False –ve	
	Negative	C: False +ve	D: True –ve	

According to the results presented in Table 2 the overall average percentage True Positive Rate (TPR) achieved by the proposed technique is 96% and the True Negative Rate (TNR) is 4% while in the same way the accuracy (AC) achieved is 98% and an Error Rate (Er.Rate) is 2%. This proves that the proposed technique is promising and robust while from computational efficiency point of view, the average processing time per slide is 1.5 seconds on Intel ® corei3-380M Processor 2.53GHz having Windows® 7 Home Basic (64-bit) operating System with 2GB Memory, showing that the proposed technique is highly efficient. However, nothing in this world is perfect the proposed technique is an attempt of improvement which achieved its goals in a simple, efficient and realistic way as compared to other techniques in the area.

Table 2 Quantitative	Analysis
----------------------	----------

Slide	Manual	Autom	TPR	AC	Er.R	TNR
No	RBCs	atic				
	Count	RBCs				
		Count				
1	31	31	1	1	0	0
2	32	32	1	1	0	0
3	31	31	1	1	0	0
4	10	8	0.8	0.89	0.11	0.2
5	20	17	0.85	0.92	0.08	0.15
6	40	39	0.98	0.99	0.01	0.03
7	40	40	1	1	0	0
8	55	54	0.98	0.99	0.01	0.02
9	77	76	0.99	0.99	0.01	0.01
10	34	33	0.97	0.99	0.01	0.03
11	88	85	0.97	0.98	0.02	0.03
12	91	90	0.99	0.99	0.01	0.01
13	21	20	0.95	0.98	0.02	0.05
14	13	13	1	1	0	0
15	15	14	0.93	0.97	0.03	0.07
16	12	11	0.92	0.96	0.04	0.08
17	10	9	0.9	0.95	0.05	0.1
18	80	79	0.99	0.99	0.01	0.01
19	66	65	0.98	0.99	0.01	0.02
20	34	33	0.97	0.99	0.01	0.03

4. CONCLUSION AND FUTURE WORK

In counting the Red Blood Cells the problems which we mentioned and solved are not only the problems of automatic studies but also they are very serious in manual studies and sometime leads to discard the slides and prepare new ones. The proposed technique achieved its goals up to its maximum level and is robust and efficient attempt in the area but still the area has capacity to improve the accuracy further. Counting of objects of interest in images is not complicated task but solving the hurdles to clear the way to reach the goal in realistic and accurate way is difficult. The consideration of the problems of clumped and overlapped Red Blood Cells will not only improve the accuracy but also reduce the burden of re-preparation of slides. The accuracy we bring to 98% still needs improvement because all health care applications are safety critical and the demand for more accuracy is always there, thus in this connection still gap existed. As future work we suggest to smoothen the boundaries will further improves the accuracy but give attention to the efficiency.

ACKNOWLEDGMENT

The first author would like to acknowledge this work to his beloved supervisor Prof.Dr.Dzulkifli Mohamad because of his love and guidance at each step makes it possible.

REFERENCES

[1] N. National Institute of Health. (1997, 27-4-2014). *Red Blood Cells Count.* Available: http://www.nlm.nih.gov/medlineplus/ency/article/0036 44.htm Last accessed on March, 2014

- [2] H. Berge, D. Taylor, S. Krishnan, and T. S. Douglas, "Improved red blood cell counting in thin blood smears," *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 204-207, 2011.
- [3] M. Buttarello and M. Plebani, "Automated blood cell counts state of the art," *American Journal of Clinical Pathology*, vol. 130, pp. 104-116, 2008.
- [4] M. G. OWAIS SHAIKH, NEHARIKA BHAT,ROSHAN SHETTY "AUTOMATED RED BLOOD CELLS COUNT," B.E. Synopsis Report, Department of Computer Engineering, Rizvi College of Engineering, University of Mumbai, Bandra(w), Mumbai - 400050, 2013-2014.
- [5] N.-T. Nguyen, A.-D. Duong, and H.-Q. Vu, "Cell splitting with high degree of overlapping in peripheral blood smear," *Int J Comp Theory Eng*, **vol**. 3, 2011.
- [6] N. H. Mahmood, P. C. Lim, S. M. Mazalan, and M. A. A. Razak, "BLOOD CELLS EXTRACTION USING COLOR BASED SEGMENTATION TECHNIQUE," *An International Journal IJLBPR*, vol.2, April 2013.
- [7] grietinfo.in.Available: grietinfo.in/projects/MAIN/BME2013/cd-8project%20report_1_.pdf Last accessed : March 2014.
- [8] N. H. Mahmood and M. A. Mansor, "Red blood cells estimation using Hough transform technique," *Signal* & *Image Processing: An International Journal (SIPIJ)*, vol. 3, pp. 53-64, 2012.
- [9] N. Ramesh, M. E. Salama, and T. Tasdizen, "Segmentation of haematopoeitic cells in bone marrow using circle detection and splitting techniques," 9th IEEE International Symposium on Biomedical Imaging (ISBI), pp. 206-209, 2012.
- [10] F. Buggenthin, C. Marr, M. Schwarzfischer, P. S. Hoppe, O. Hilsenbeck, T. Schroeder, and F. J. Theis, "An automatic method for robust and fast cell detection in bright field images from high-throughput microscopy," *BMC bioinformatics*, vol. 14, p. 297, 2013.
- [11] K. Prasad, J. Winter, U. M. Bhat, R. V. Acharya, and G. K. Prabhu, "Image analysis approach for development of a decision support system for detection of malaria parasites in thin blood smear images," *Journal of digital imaging*, vol. 25, pp. 542-549, 2012.
- [12] P. A. C. Amit Kumar, Prof. P. U. Tembhare, Prof. C. R. Pote, "Enhanced Identification of Malarial Infected Objects using Otsu Algorithm from Thin Smear Digital Images," *International Journal of Latest Research in Science and Technology*, vol.1, pp. Page No159-163 August(2012) 2012.
- [13] A. LaTorre, L. Alonso-Nanclares, S. Muelas, J. Peña, and J. DeFelipe, "Segmentation of neuronal nuclei based on clump splitting and a two-step binarization of images," *Journal of Expert Systems with Applications*, vol. 40, pp. 6521-6530, 2013.
- [14] S. Tafavogh, K. F. Navarro, D. R. Catchpoole, and P. J. Kennedy, "Segmenting Neuroblastoma Tumor Images and Splitting Overlapping Cells Using Shortest Paths

between Cell Contour Convex Regions," *ed: Springer* of Artificial Intelligence in Medicine, pp. 171-175, 2013.

- [15] C. Zhang, C. Sun, R. Su, and T. D. Pham, "Segmentation of clustered nuclei based on curvature weighting," in *Proceedings of the 27th Conference on Image and Vision Computing New Zealand*, pp. 49-54, 2012.
- [16] H. Wang, H. Zhang, and N. Ray, "Clump splitting via bottleneck detection," 18th IEEE International Conference on Image Processing (ICIP), pp. 61-64, 2011.
- [17] S. K. Kumarasamy, S. Ong, and K. S. Tan, "Robust contour reconstruction of red blood cells and parasites in the automated identification of the stages of malarial infection," *Journal of Machine Vision and Applications*, vol. 22, pp. 461-469, 2011.
- [18] Q. Wen, H. Chang, and B. Parvin, "A Delaunay triangulation approach for segmenting clumps of nuclei," *IEEE International Symposium on Biomedical Imaging: From Nano to Macro ISBI*. pp. 9-12, 2009.
- [19] V. V. Makkapati and S. K. Naik, "Clump splitting based on detection of dominant points from contours," in Automation Science and Engineering, 2009. IEEE International Conference on CASE, pp. 197-201, 2009.
- [20] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological image analysis: A review," *IEEE Reviews in Biomedical Engineering*, , vol. 2, pp. 147-171, 2009.
- [21] F. Cloppet and A. Boucher, "Segmentation of overlapping/aggregating nuclei cells in biological images," 19th International Conference on Pattern Recognition, ICPR, pp. 1-4, 2009.
- [22] H. Tulsani, "Segmentation using morphological watershed transformation for counting blood cells," *IJCAIT*, vol. 2, pp. 28-36, 2013.
- [23] L. Ferro, P. Leal, M. Marques, J. Maciel, M. I. Oliveira, M. A. Barbosa, and P. Quelhas, "Multinuclear Cell Analysis Using Laplacian of Gaussian and Delaunay Graphs," in *Pattern Recognition and Image Analysis*, ed: Springer, pp. 441-449, 2013.
- [24] E. Hodneland, T. Kögel, D. M. Frei, H.-H. Gerdes, and A. Lundervold, "CellSegm-a MATLAB toolbox for high-throughput 3D cell segmentation," *Source code for biology and medicine*, vol. 8, pp. 1-24, 2013.
- [25] O. Schmitt and S. Reetz, "On the decomposition of cell clusters," *Journal of Mathematical Imaging and Vision*, vol. 33, pp. 85-103, 2009.
- [26] O. Schmitt and M. Hasse, "Morphological multiscale decomposition of connected regions with emphasis on cell clusters," *Computer Vision and Image Understanding*, vol. 113, pp. 188-201, 2009.
- [27] V. Špringl, "Automatic Malaria Diagnosis through Microscopy Imaging," *Higher Diploma, FACULTY OF*

ELECTRICAL ENGINEERING, CZECH TECHNICAL UNIVERSITY IN PRAGUE, PRAGUE, 2009.

- [28] W. N. Gonçalves and O. M. Bruno, "Automatic system for counting cells with elliptical shape," *arXiv preprint arXiv:1201.3109*, 2012.
- [29] H. Kong, M. Gurcan, and K. Belkacem-Boussaid, "Splitting touching-cell clusters on histopathological images," in 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 208-211, 2011.
- [30] H. Kong, M. Gurcan, and K. Belkacem-Boussaid, "Partitioning histopathological images: an integrated framework for supervised color-texture segmentation and cell splitting," *IEEE Transactions on Medical Imaging*, vol. 30, pp. 1661-1677, 2011.
- [31] M. Köppen, K. Yoshida, and P. A. Valle, "Gestalt theory in image processing: A discussion paper," in Proceedings of three rivers on Soft Computing in Industrial Applications, University of Passau, Germany, pp. 1-3, August 1 - 3, 2007.
- [32] H. Jiang, C.-W. Ngo, and H.-K. Tan, "Gestalt-based feature similarity measure in trademark database," *Jouranl of Pattern recognition*, vol. 39, pp. 988-1001, 2006.
- [33] U. Airsang, V. Ghorpade, and S. T. Rajamani, "Contour feature-point tagging as a mechanism for automated splitting of highly-occluded and dissimilarsized cells in blood smear images," *IEEE Second International Conference on Image information Processing (ICIIP)*, pp. 470-475, 2013.
- [34] DPDx, "DPDx, Laboratory Identification of Parasites, Centers of diseases Control and Prevention," D. a. m. b. C. s. D. o. P. D. a. M. (DPDM), Ed., ed. 1600 Clifton Rd. Atlanta, GA 30333, USA, Web site developed and maintained by CDC's Division of Parasitic Diseases and Malaria (DPDM), http://www.dpd.cdc.gov/dpdx/default.htm 2002. Last accessed March.2014.
- [35] N. Abbas and D. Mohamad, "Automatic Color Nuclei Segmentation of Leukocytes for Acute Leukemia," *Research Journal of Applied Sciences, Engineering and Technology*, **vol.** 7(14), pp. 2987-2993, 2014.
- [36] N. ABBAS and D. MOHAMAD, "MICROSCOPIC RGB COLOR IMAGES ENHANCEMENT FOR BLOOD CELLS SEGMENTATION IN YCbCr COLOR SPACE FOR K-MEANS CLUSTERING," Journal of Theoretical & Applied Information Technology, vol. 54, pp. 72-78, 2013.