

Semi-Automatic Red Blood Cells Counting in Microscopic Digital Images

Naveed Abbas*, Dzulkifli Mohamad, Abdul Hanan Abdullah

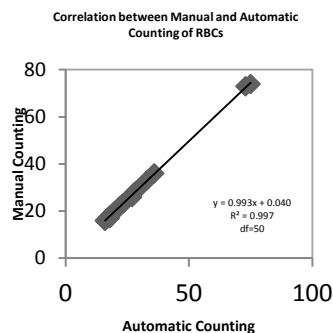
ViCube Research Lab, Department of Software Engineering, Faculty of Computing, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia

*Corresponding author: naveed23a@yahoo.com

Article history

Received :10 December 2014
Received in revised form :
1 February 2015
Accepted :12 February 2015

Graphical abstract



Abstract

The main purpose of this study is to employ the modern technologies and techniques to semi-automate the quantification process of the Red Blood Cells in Microscopic thin blood smear digital images. The process needs to be more accurate, efficient and universal than the currently practiced methods. The study considers the process to be semi-automated for two reasons, i.e. due to the critical aspect of life and due to the diverse nature of the Red Blood Cells in cluster formation. The methodology of this study involved interactive simple cuts and morphological operations for splitting clusters of Red Blood Cells while counting is carried out through a labeling matrix. The Red Blood Cells counting is part of the complete blood count test and is frequently suggested by the physician to know the number of Red Blood Cells in the patient's body. The proposed method considers the counting process of the Red Blood Cells: first split the clusters and then count the Red Blood Cells. The proposed method achieved an overall True Positive Rate (TPR) of 0.997%, True Negative Rate (TNR) of 0.00265%, accuracy of 0.998% and average error rate of 0.001375% tested on 50 images. Data set also on the same number of images linear correlation coefficient R^2 is 0.997 between manual and semi-automatic counting of Red Blood Cells.

Keywords: Clustered red blood cells; healthcare applications; complete blood test; occlusions; rouleaux splitting

© 2015 Penerbit UTM Press. All rights reserved.

1.0 INTRODUCTION

The counting process of Red Blood Cells is demanding in various blood tests because the alteration of number of Red Blood Cells from normal range in both cases (Low and High) is a sensitive indicator about serious disorder in the body. The number of Red Blood Cells high then normal range indicates Kidney tumor, Heart diseases, Low Blood oxygen level etc. 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 hematologists increased the demand of automatic and semi-automatic counting techniques to overcome 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 a difficult task but for high accuracy it involves several other problems i.e. image pre-processing, splitting of clustered Red Blood Cells. If these mentioned problems are not addressed in proper way then the

accuracy will be on the stack because the clustered Red Blood Cells are appeared as a single area and in reality it is a combination of multiple Red Blood Cells. Further, the clustered are divided into Clumped and Overlapped Red Blood Cells. Clumps of Red Blood Cells occurred in the case when iron deficiency exists in the blood, the Red Blood Cells glued each other and formed long chains while overlapped Red Blood Cells are formed due to improper slide preparation and is also considered as a big problem in the manual microscopy because it leads to discard the slide and prepare another one. This study considered all these problems and after solving the given problems, then count the Red Blood Cells.

Recently, too many efforts 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. This study considers the proper and accurate solution to be interactive means semi-automatic because previously we proposed six novel automatic techniques on different grounds and show high success rates but there are situations in which the Red Blood Cells are clumped and overlapped densely, splitting of which is impossible by any automatic method and this is the main reason that we switched from automatic to semi-automatic to involve human

expertise. Moreover, in interactive cuts based strategy we struggled to involve minimum intervention. 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. 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, 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]. The problem in this approach seems to be un-realistic as due to its highly complex nature and implementation.

2.0 EXPERIMENTAL

In this paper we performed the experimentations on a set of Microscopic thin blood smear digital images, which were obtained from the [32], freely available for research purposes. The proposed methodology started with image pre-processing then the slide image is checked for clustered Red Blood Cells if existed then passed from interactive splitting process of Red Blood Cells clusters because without splitting the accuracy is

compromised on the other hand if clustered Red Blood Cells not existed then the control is directly transferred to counting the Red Blood Cells. This whole process is presented as overall methodology of this study and the simulated diagram of the whole process in the form of images are depicted in (Figures 1 and 2).

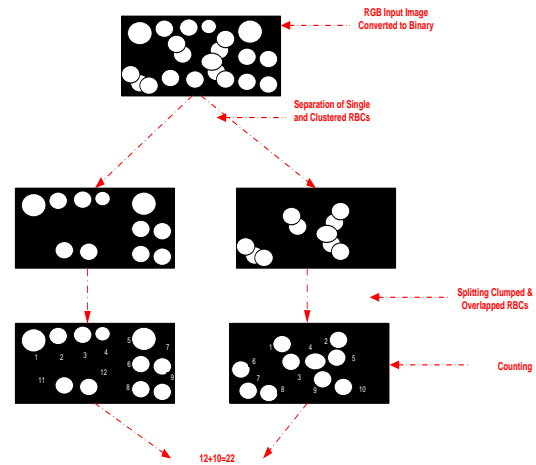


Figure 1 Simulated images of the whole process

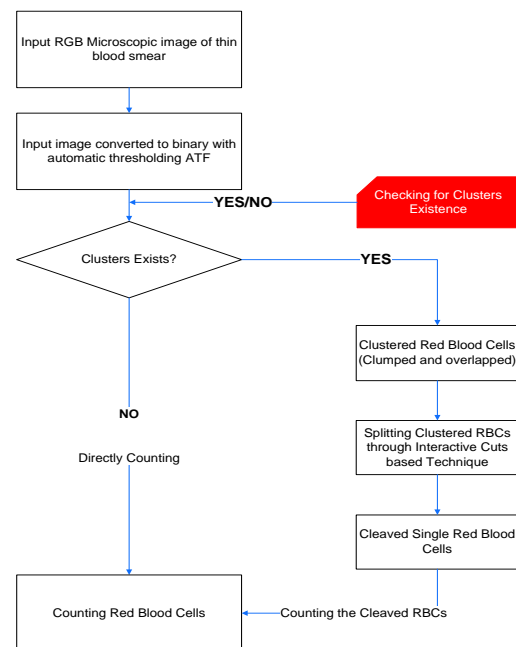


Figure 2 Overall methodology

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. After conversion small areas are identified as noise and removed from the binary image and holes in the centers of the RBCs, formed due to hemoglobin in the centers of the RBCs and its similarity to the background are filled and we get the image presented in Figure 3 which is ready for further processing.

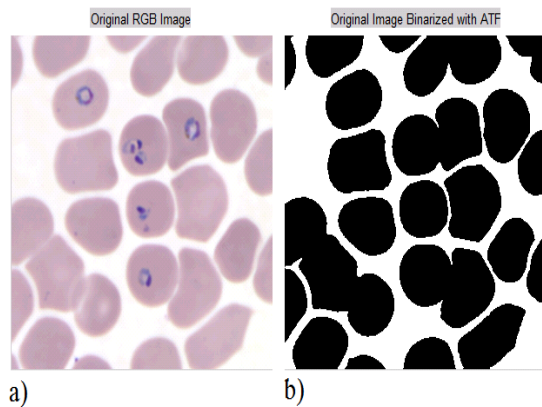


Figure 3 Matlab results a) Original RGB image b) Pre-processed image

To check the existence of clustered of RBCs; we applied a double check on all the RBCs. We find the areas and elongation of the convex hulls of the RBCs as mentioned in Equation 1 and 2 respectively. Next between these two measures we find normalize variance among all the RBCs and empirically through much experimentation, we found that if the variance is high 0.2 in case of area and high than 0.5 in case of elongation will be considered as clustered RBCs existed as mentioned in Equation 3.

$$\text{Area} = \text{No. of Pixels}; \quad (1)$$

where, No.of Pixels= Pixels in the convex hull object

$$\text{Elongation} = \frac{\text{Length}}{\text{Breadth}} \quad (2)$$

where, Length =Major Axis and Breadth = Minor Axis

$$\sigma^2 = \frac{(X - \mu)^2}{N} \quad (3)$$

where, X represents the area or elongation, N is the no. of terms in distribution.(3)

After confirmation that clustered Red Blood Cells existed, then this step is eligible for execution, otherwise the control is directly transferred to the counting process. The splitting starts by loading the grayscale image of the original RGB image for the purpose of clear visualization along with the pointer automatically for interactive cuts. The cuts made by the experts are in the form of points having possible gaps. For accuracy purpose we find the distance between the successive points with the equation mentioned as 4. The gaps among the points are filled by adding 0.5 with the distance between the successive points to form smooth cuts. The smooth cuts are then mapped in the binary image with a slight uniform erosion to split the clustered Red Blood Cells into cleaved single Red Blood Cells. In this whole process only the cuts are on the part of expert the rest whole process is automatic.

$$\text{Distance, } D = \sqrt{\frac{(x_2 + x_1)^2 - (y_2 + y_1)^2}{2}} \quad (4)$$

Filling Gaps = D+0.5 for all the points in cuts

The whole concept is diagrammatically simulated in Figure 5 while the results obtained from Matlab are presented as a, b, c, d and e in (Figure 4).

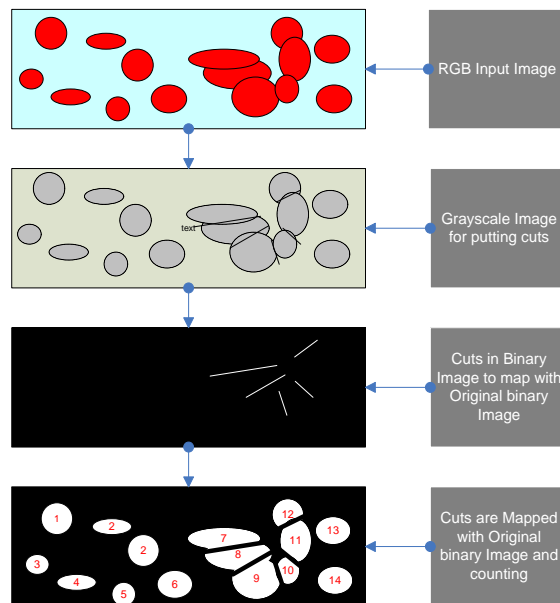


Figure 4 Simulated Image showing the concept behind the whole process

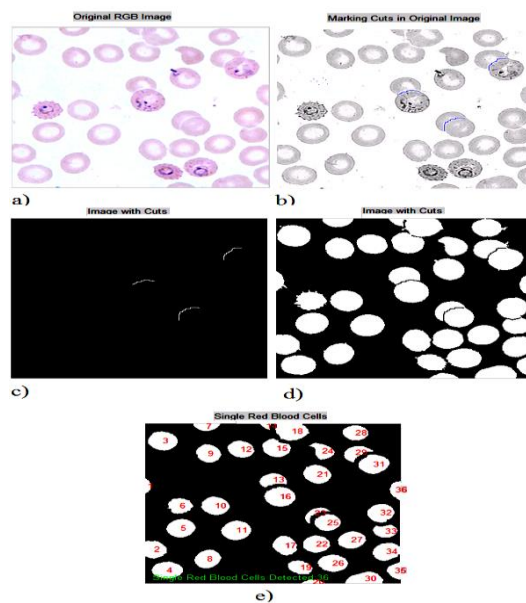


Figure 5 Matlab results a) Presents original RGB image, b) Presents grayscale image for putting cuts, c) Presents the binary image of the cuts made interactively, d) Presents the image in which cuts are mapped to the original binary image, e) Presents the slightly eroded image at cuts to split the clusters into single red blood cells and also counted

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.

3.0 RESULTS AND DISCUSSION

In this section we analyzed the results through visual inspection and statistically on microscopic thin blood smear digital images dataset of 50 images obtained from [32]. In visual inspection based analysis we divide the images into two groups i.e. images having clustered Red Blood Cells and Images without clustered Red Blood Cells for presentation through visual inspection with ground reality. In (Figure 6) on the basis of two input images we presented that what are the effects on the accuracy when the clustered RBCs are not addressed.

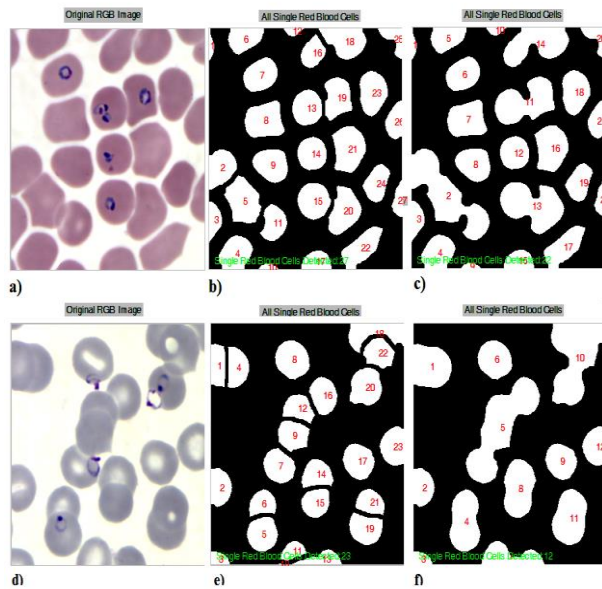


Figure 6 a) and d) Present original RGB image with clustered RBCs while b) and e) Present the binary image having numbers of each RBC and the total number of RBCs is shown in the green color when clustered RBCs are cleaved and finally c) and f) present counting without splitting the clustered RBCs

According to (Figure 6), we can clearly know about the accuracy that without addressing the clustered RBCs splitting the number of RBCs in the c) are 22 while after splitting it becomes 27 as shown in b), also the same situation we faced in e) the actual number is 23 while without splitting it is 12. The rest of the results as shown in (Figures 7 and 8) are on images having no clustered RBCs.

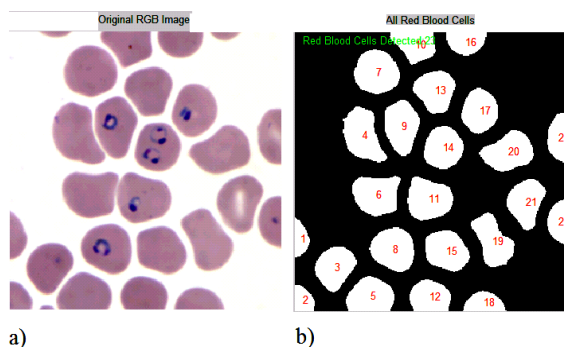


Figure 7 a) Presents original RGB image without clustered RBCs while b) Presents the binary image having numbers of each RBC and the total number of RBCs is shown in the green color

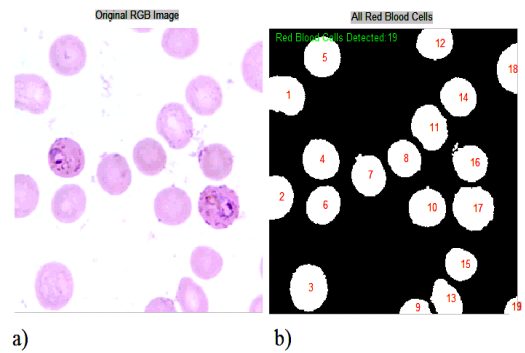


Figure 8a) Presents original RGB image without clustered RBCs while b) Presents the binary image having numbers of each RBC and the total number of RBCs is shown in the green color

According to (Figures 7 and 8), the number of RBCs has no problem because clustered RBCs do not exist but the situation as shown in (Figure 9) having two images with clustered Red Blood Cells have problems, thus after splitting the problems are resolved.

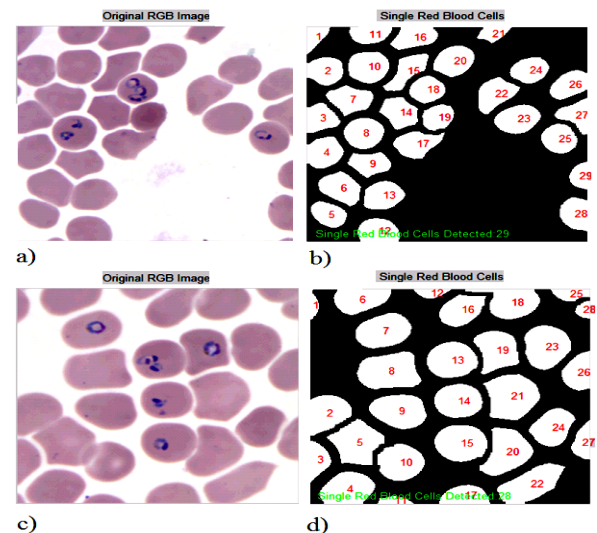


Figure 9 a) and c) Present original input RGB images of thin blood smears while images b) and d), are the final output images after applying the proposed methods in this paper

Analysis through statistical metrics for measuring the accuracy the experimentation is performed on the images data set of 50 images and compared the results of semi-automatically counted Red Blood Cells (by computer) with the manually counted Red Blood Cells (by medical experts) through confusion matrix presented as Table 1, to 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 5, 6, 7 and 8 respectively.

$$TPR = \frac{A}{A+B} \quad , \quad AC = \frac{A+D}{A+B+C+D} \quad (5 \text{ and } 6)$$

$$Er. R = 1 - AC \quad , \quad TNR = \frac{D}{C+D} \quad (7 \text{ and } 8)$$

Table 1 Confusion matrix

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

Confusion Matrix Based Analysis

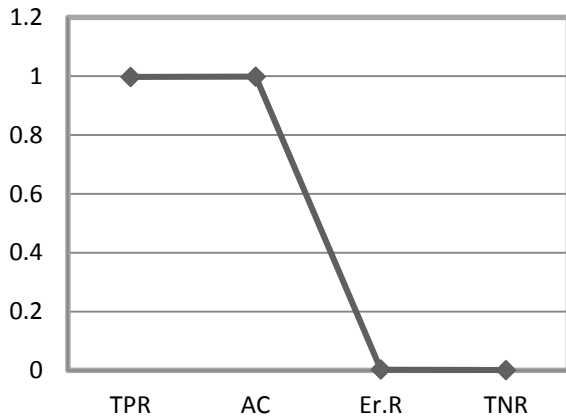


Figure 10 Confusion matrix based analysis

The proposed technique is experimented on 50 images and the TPR, Accuracy, Error rate and TNR are presented graphically in (Figure 10). Further, we also performed the linear correlation analysis to know the relationship through Pearson's Correlation presented graphically in (Figure 11).

Correlation between Manual and Automatic Counting of RBCs

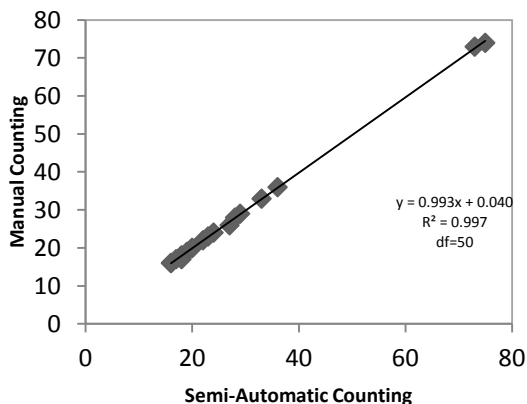


Figure 11 Correlation between manual and semi-automatic counting of red blood cells

According to the graph presented in (Figure 11), the correlation coefficient R^2 shows strong positive linear correlation between the manual and semi-automatic counting of Red Blood Cells, that indicates a reliable relationship between the manual and semi-automatic counting.

4.0 CONCLUSION

The main and challenging job in counting Red Blood Cells is that of clustered Red Blood Cells splitting. Splitting clusters of Red Blood Cells plays a vital role in improving the accuracy. The interactive cuts based solution is very simple and efficient way of involving minimum human intervention. The proposed method achieved an overall sensitivity of 0.997%, while accuracy 0.998% and in the same way the achieved overall True Negative Rate is 0.00265% while Error Rate is 0.001375%. The proposed method also shows a strong relationship with manual counting as the correlation coefficient, $R^2=0.997$ All these are highly achieved statistics in the area.

Acknowledgement

The authors would like to pay thanks to Prof.Dr.IkramulMabood (Medical expert), Prof.Dr. AmreekLal, (Pathologist), Swat Medical College, Saidu Sharif Swat, KPK, Pakistan and Dr.Ikram ur Rehman, (Medical Expert), regional coordinator NSRSP, swat, KPK, Pakistan for their valuable guidance in the Medical field and manual counting of Red Blood Cells in thin blood smear digital images.

References

- [1] National Institute of Health, N. 1997. *Red Blood Cells Count*. Retrieved on 10-4-2014 from: <http://www.nlm.nih.gov/medlineplus/ency/article/003644.htm>.
- [2] Berge, H., Taylor, D., Krishnan, S. and Douglas, T. S. 2011. Improved Red Blood Cell Counting in Thin Blood Smears. *Proceedings of the 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. 30 March–2 April, McCromok Palace, Chicago: IEEE. 204–207.
- [3] Buttarello, M. and M. Plebani, 2008. Automated Blood Cell Counts State of the Art. *American Journal of Clinical Pathology*. 130(1):104–116.
- [4] Shaikh, M. G. O., Bhat, N. and Shetty, R. 2013–2014. *Automated Red Blood Cells Count*. B.E., University of Mumbai, Bandra, Mumbai.
- [5] Nguyen, N.-T., Duong, A.-D. and Vu, H.-Q. 2011. Cell Splitting with High Degree of Overlapping in Peripheral Blood Smear. *Int J Comp Theory Eng*. 3(3).
- [6] Mahmood, N. H., Lim, P. C., Mazalan, S. M. and Razak, M. A. A. 2013. Blood Cells Extraction Using Color Based Segmentation Technique. *Int. J. LifeSc. Bt & Pharm. Res*. 2(2): 233–240.
- [7] Grietinfo. 2013. Detection of Abnormal Red Blood Cells Using Matlab. Retrieved 20 March 2013 from grietoinfo.in/projects/Main/BME2013/cd-8.
- [8] Mahmood, N. H. and Mansor, M. A. 2012. Red Blood Cells Estimation Using Hough Transform Technique. *Signal & Image Processing: An International Journal (SIPIJ)*. 3(2): 53–64.
- [9] Ramesh, R., Salama, N., Tasdizen, M. E. and Tolga, T. 2012. Segmentation of Haematopoietic Cells in Bone Marrow Using Circle Detection and Splitting Techniques. *Proceedings of the 2012 IEEE 9th International Symposium on Biomedical Imaging (ISBI)*. 2–5 May, Barcelona, Spain, IEEE. 206–209.
- [10] Buggenthin, F., Marr, C., Schwarzfischer, M., Hoppe, P. S., Hilsenbeck, O., Schroeder, T. and Theis, F. J. 2013. An Automatic Method for Robust and Fast Cell Detection in Bright Field Images from High-throughput Microscopy. *BMC bioinformatics*. 14(1): 297.
- [11] Prasad, K., Winter, J., Bhat, U. M., Acharya, R. V. and Prabhu, G. K. 2012. Image Analysis Approach for Development of a Decision Support System for Detection of Malaria Parasites in Thin Blood Smear Images. *Journal of Digital Imaging*. 25(4): 542–549.
- [12] Kumar, A., Choudhary, A., Tembhare, P.U. and Pote, C. R. 2012. Enhanced Identification of Malarial Infected Objects using Otsu Algorithm from Thin Smear Digital Images. *International Journal of Latest Research in Science and Technology*. 1(2): 159–163.
- [13] Latorre, A., Alonso-Nanclares, L., Muelas, S., Peña, J. and DeFelipe, J. 2013. Segmentation of Neuronal Nuclei Based on Clump Splitting and

- A Two-step Binarization of Images. *Expert Systems with Applications*. 40(16): 6521–6530.
- [14] Tafavogh, S., Navarro, K. F., Catchpole, D. R. and Kennedy, P. J. 2013. Segmenting Neuroblastoma Tumor Images and Splitting Overlapping Cells Using Shortest Paths between Cell Contour Convex Regions *Artificial Intelligence in Medicine*. Springer. 171–175.
- [15] Zhang, C., Sun, C., Su, R. and Pham, T. D. 2012. Segmentation of Clustered Nuclei Based on Curvature Weighting. *Proceedings of the 2012 IEEE 27th Conference on Image and Vision Computing*. 27–29 November. New Zealand: ACM. 49–54.
- [16] Wang, H., Zhang, H. and Ray, N. 2011. Clump Splitting Via Bottleneck Detection. *Proceedings of the 2011 IEEE 18th International Conference on Image Processing (ICIP)*, 11–14 September, Bursel, Belgium: IEEE. 61–64.
- [17] Kumarasamy, S. K., Ong, S. and Tan, K. S. 2011. Robust Contour Reconstruction of Red Blood Cells and Parasites in the Automated Identification of the Stages of Malarial Infection. *Machine Vision and Applications*. 22(3): 461–469.
- [18] Wen, Q., Chang, H. and Parvin, B. 2009. A Delaunay Triangulation Approach for Segmenting Clumps of Nuclei. *Proceedings of the 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. ISBI*. 28 June–1 July. Boston, MA, USA: IEEE. 9–12.
- [19] Makkapati, V. V. and Naik, S. K. 2009. Clump Splitting Based on Detection of Dominant Points from Contours. *Proceedings of the 2009 IEEE International Conference on Automation Science and Engineering*, 22–25 August, Bangalore, India: IEEE. 197–201.
- [20] Gurcan, M. N., Boucheron, L. E., Can, A., Madabhushi, A., Rajpoot, N. M. and Yener, B. 2009. Histopathological Image Analysis: A Review. *IEEE Reviews in Biomedical Engineering*. 2: 147–171.
- [21] Cloppet, F. and Boucher, A. 2008. Segmentation of Overlapping/Aggregating Nuclei Cells in Biological Images. *Proceedings of the 2008 IEEE 19th International Conference on Pattern Recognition, ICPR*, 8–11 December, Florida, USA: IEEE. 1–4.
- [22] Tulsani, H. 2013. Segmentation Using Morphological Watershed Transformation for Counting Blood Cells. *IJCAIT*. 2(3): 28–36.
- [23] Ferro, L., Leal, P., Marques, M., Maciel, J., Oliveira, M. I., Barbosa, M. A. and Quelhas, P. 2013. Multinuclear Cell Analysis Using Laplacian of Gaussian and Delaunay Graphs, In Sanches, Joao Miguel, Micó, Luisa, Cardoso, Jaime (Eds.). *Pattern Recognition and Image Analysis*. Oslo Norway: Springer Berlin Heidelberg. 441–449.
- [24] Hodneland, E., Kögel, T., Frei, D. M., Gerdes, H.-H. and Lundervold, A. 2013. CellSegm-a MATLAB Toolbox for High-Throughput 3D Cell Segmentation. *Source Code for Biology and Medicine*. 8(1): 1–24.
- [25] Schmitt, O. and Reetz, S. 2009. On the Decomposition of Cell Clusters. *Journal of Mathematical Imaging and Vision*. 33(1): 85–103.
- [26] Schmitt, O. and Hasse, M. 2009. Morphological Multiscale Decomposition of Connected Regions with Emphasis on Cell Clusters. *Computer Vision and Image Understanding*. 113(2): 188–201
- [27] Springl, V. 2009. *Automatic Malaria Diagnosis through Microscopy Imaging*. Higher Diploma, Czech Technical University In Prague, Prague.
- [28] Goncalves, W. N. and Bruno, O. M. 2012. Automatic System for Counting Cells with Elliptical Shape. *Cornell University Library, arXiv:1201.3109*. 9(1).
- [29] Kong, K., Gurcan, H., Boussaid, B. M. and Kamel, K. 2011b. Splitting Touching-cell Clusters on Histopathological Images. *Proceedings of the 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 30 March–2 April, Chicago, USA: IEEE. 208–211.
- [30] Kong, K., Gurcan, H., Boussaid, B. M. and Kamel, K. 2011a. Partitioning Histopathological Images: An Integrated Framework for Supervised Color-texture Segmentation and Cell Splitting. *IEEE Transactions on Medical Imaging*. 30(9): 1661–1677.
- [31] Koppen, K., Yoshida, M., Valle, P. and Pablo, A. 2007. Gestalt Theory in Image Processing: A Discussion Paper. *Proceedings of the 2007 Three Rivers workshop on Soft Computing in Industrial Applications*, 1–3 August, University of Passau, Germany. 1–5.
- [32] DPDx. 2002. DPDx, Laboratory Identification of Parasites, *Centers of diseases Control and Prevention, Division of Parasitic Diseases and Malaria (DPDM)* Govt of USA. Retrieved 15 March 2013, from <http://www.dpd.cdc.gov/dpdx/default.htm>.