Journal of Engineering and Applied Sciences 12 (Special Issue 12): 9496-9501, 2017
ISSN: 1816-949X

Segmentation and Clumped Cell Detection in Microscopic Peripheral Blood Smear Images

[1]Dhanya Bibin and [2]P. Punitha
[1]Department of Research and Development Centre, Bharathiar University, 641046 Coimbatore, India
[2]Department of MCA, PES Institute of Technology, 560085 Bangalore, India

Abstract: This study proposes an efficient approach for automatic detection of clumped cells in peripheral blood smear images that combines Chan Vese segmentation and ellipse fitting. Clumped cell detection is a critical task in developing efficient blood cell segmentation methods in automated malaria diagnosis systems. One of the major challenges in the segmentation of blood cell images is the presence of closely clumped cells. The proposed clumped cell detection method has three steps. Initially, the image segmentation is performed using Chan Vese algorithm to extract clumped and non clumped blood cells from the background. Then, two post-processing techniques are applied on the segmented image to eliminate artefacts, platelets and holes which are irrelevant in this study. The artefacts and platelets are removed using a size based thresholding method and hole elimination is achieved using morphological reconstruction. Then, a robust and efficient ellipse fitting is performed on the connected components in the segmented image to differentiate them as clumped cells or non-clumped cells. The differentiation is based on a threshold value which is computed from the eccentricity of the fitted ellipse. The robustness and efficiency of the proposed method has been evaluated by comparing empirically to manual detection performed by haematologists.

Key words: Malaria diagnosis, clump cell detection, ellipse fitting, eccentricity, Chan Vese algorithm, peripheral blood smear

INTRODUCTION

Malaria is a life-threatening disease caused by plasmodium parasites. According to WHO, there were about 400 million cases of malaria and an estimated 627 K deaths. Delay in diagnosis and treatment is the leading cause of death in malaria patients. It is due to the lack of experienced medical professionals in rural areas (WHO., 2017). Even though new techniques have emerged for the diagnosis of malaria, manual microscopy examination of peripheral blood smears is the "the gold standard" and most prevalent diagnostic technique for malaria detection (WHO., 2017; World Health Organization, 1991). In microscopy diagnosis, the pathologist observes the morphological characteristics of the blood cell components in a blood slide to identify the infected blood cells. This process is time-consuming, laborious and leads to fatigue, especially, in peak infection seasons and can lead to wrong diagnosis and treatment which could even result in death of patients. Quality and accuracy of the manual microscopy method depends on the skill and experience of the pathologist. So, manual microscopy is not considered as a reliable screening method for malaria diagnosis (World Health Organization, 1991). Hence, there is a need to automate the microscopy malaria diagnosis process. Automation of malaria microscopy diagnosis is a larger problem involving various critical stages.

In this study, we address the problem of clumped/overlapping cell detection which is the basis for any further analysis stages in malaria diagnosis process. The detection and splitting of clumped cells into individual cells is very important to identify the parasite infected cells and to estimate the infection severity to determine correct medication to the patient. We propose a two stage algorithm for clumped blood cell detection which combines Chan Vese segmentation algorithm and ellipse fitting with two post-processing techniques to improve the detection accuracy.

Related research in blood cell segmentation and clump cell splitting: Recently, researchers have made considerable amount of studies related cell image segmentation (Ruberto *et al*., 2001, 2002; Mao *et al*., 2006; Yang *et al*., 2006; Ritter and Cooper, 2007; Kumar *et al*.,

Corresponding Author: Dhanya Bibin, Department of Research and Development Centre, Bharathiar University, 641064 Coimbatore, India

2006; Diaz *et al*., 2007, 2009; Cheng and Rajapakse, 2009; Tek *et al*., 2005; Zhou and Mao, 2006; Kong *et al*., 2011; Wang *et al*., 2012) and clumped cell splitting. The performance of any automated cell image analysis systems largely dependent on the efficiency of the cell segmentation and clump cell splitting algorithm. Clump splitting methods are available in literature based on morphological operation (Ruberto *et al*., 2001, 2002); watershed techniques (Mao *et al*., 2006; Yang *et al*., 2006); concavity analysis (Ritter and Cooper, 2007; Kumar *et al*., 2006) and model based approaches (Diaz *et al*., 2007, 2009). The concavity based methods are very sensitive in fixing the threshold to identify the concave points and watershed techniques are time consuming and prone to over segmentation. To overcome these problems and to obtain efficient segmentation, region merging and marker-controlled watershed techniques have been introduced (Cheng and Rajapakse, 2009; Tek *et al*., 2005). Morphology based clump cell splitting methods are too sensitive to the selection of structuring elements and thresholds (Yang *et al*., 2006; Zhou and Mao, 2006). The focus of the above mentioned researches are primarily around splitting of clumped cells. However in all of the above researches, the researchers have treated the clumped and non-clumped cells uniformly. But in a well treated blood smear image, the percentage of clumped cells would be considerably less than the non-clumped cells and hence, treating only the clumped cells will provide higher efficiency and computational performance for the algorithms. Very few researchers have attempted to detect the clumped cells before proceeding to the clump splitting algorithm (Kong *et al*., 2011; Wang *et al*., 2012). Kong *et al*. (2011) and Wang *et al*. (2012) used a shape classifier to determine whether a connected component is clumped or non clumped. Wang *et al*. (2012), the differentiation of clumped and non-clumped cells is based on the distance between the most likely radial symmetry center and the geometrical center of the connected component. Since, the above mentioned researches are not specifically for blood smear images and only shape or size based clump cell detection techniques are not sufficient to identify the clumped cells in this particular application in this study, we propose a new methodology to distinguish the clumped and single blood cells.

MATERIALS AND METHODS

Proposed method: Our proposed clumped cell detection algorithm has 2 stages. In Stage 1, we segment the blood smear image using Chan Vese algorithm (Chan and Vese, 2001). This stage eliminates the background and extracts the foreground regions such as RBCs, WBCs, platelets and artifacts. Two post-processing operations are performed on the segmented image to eliminate the insignificant components such as artifacts, platelets and holes present in the middle of the RBCs. In a blood smear image, artifacts represent bacteria, spores, crystallized stain chemicals and dirt whereas platelets are small, irregularly shaped bodies. Peripheral blood smear images contain large number of clumped RBC and WBC cells and the separation of these cells into individual cells is very important in the computerized diagnosis of malaria. In blood smear images, single RBCs and WBCs are almost circular in shape whereas the clumped cells are typically elliptical. Parasite infected RBCs and WBCs are typically larger in size as compared to the healthy RBC/WBCs, though they retain their circular shape. One may feel that the area of cells can help in detecting the clumped cells but this is not true as the infected RBCs and WBCs are typically larger in size and will be falsely identified as clumped cell. Since, the clumped cells are of the shape of a squashed circle in Stage 2, the clump cell detection is achieved by fitting the best possible ellipse on the contour of the blobs of the segmented image and separate the blobs as clumped and non clumped cells based on a threshold value computed from the eccentricity of the ellipse. A schematic overview of the proposed blood smear image segmentation and clumped cell detection algorithm is provided in Fig. 1.

Segmentation: This research uses Chan Vese algorithm (Chan and Vese, 2001) to extract the foreground blood cell components of the blood smear image. Peripheral blood smear images have difficult topology such as disconnected regions, multiple holes, blurred edges and overlapping cell structures. They are also nonuniformly illuminated due to the usage of different kinds of cameras. Active contour models are ideal for the segmentation of images having such complex structures. Chan and Vese (2001), proposed an energy minimization of the image to detect edges of objects embedded within an image. This model begins with a contour in the image plane defining an initial segmentation and evolves this contour in such a way that it stops on the boundaries of the foreground region. The Chan Vese Model is the curve evolution implementation of a piecewise-constant case of the Mumford-Shah Model (Mumford and Shah, 1989). In this region based method, the image energy is computed from the intensity variances inside and outside of the contour. The resultant image obtained after segmentation contains RBCs, WBCs, parasites, artifacts and platelets is shown in Fig. 2b.

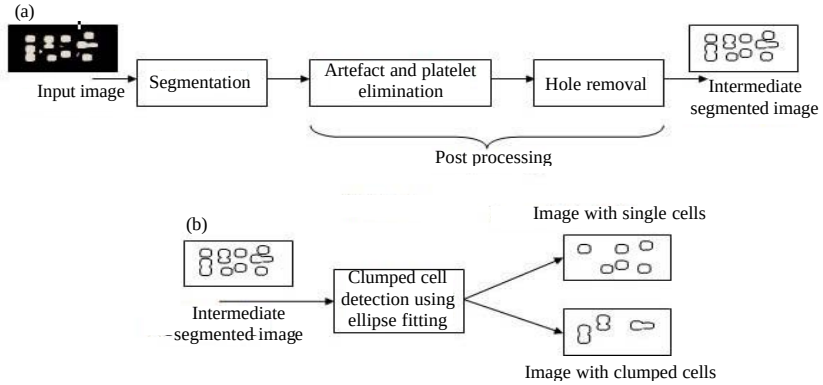

Fig. 1: Schematic overview of the proposed algorithm: a) Stage 1 and b) Stage 2

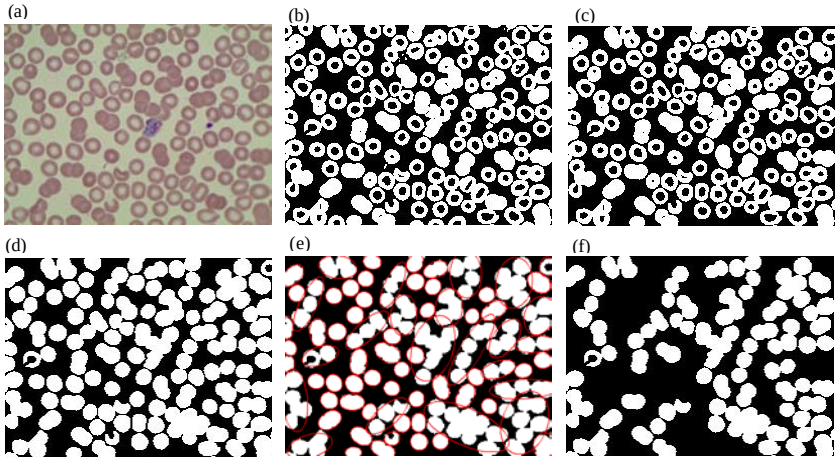

Fig. 2: Illustration of the clumped cell detection algorithm: a) Original blood smear image; b) Segmented image; c) Artifacts and platelets removed image; d) Hole removed image; e) Image with ellipse fitted on the blood cells and f) Image with clumped cells

Postprocessing: The post processing stage involves elimination of the artifacts and platelets, followed by hole filling as the presence of artifacts, platelets and holes lead to inaccuracy in the subsequent clump cell detection and clump splitting operation.

Artefact and platelet elimination: Since, the size of artifacts and platelets are comparatively smaller than the RBCs and WBCs, this research empirically decided an optimum area threshold to eliminate them from the segmented image. To find the optimum threshold, firstly the segmented image is subjected to connected component labeling (Wang *et al*., 2012). The area of each connected component (RBCs, WBCs, platelets and artifacts) is computed by counting the number of pixels in them. Empirically, it has been observed that the difference in the size of artifacts and platelets from RBCs and WBCs in terms of number of pixels is in between 115-120 pixels. We have chosen 117 as the optimum threshold. The proposed approach eliminated majority of the artefacts and platelets. The resultant image of this procedure is shown in Fig. 2c.

Hole removal: A procedure based on morphological reconstruction is utilized for hole removal. A hole in a binary image is a set of background pixels enclosed by foreground pixels. In a segmented image, the background of the image is connected to the edges of all foreground objects. We have filled the holes using a flood-fill operation starting with the border pixels of the background to fill the background area connected to the edges of the foreground regions and identify the holes in the image as background pixels that cannot be reached by such operation. Such holes can be filled by inverting the values of the pixels that were not reached by the flood-fill operation and mark them as foreground. The hole filling procedure is carried out as follows:

Let I(x,y) denote a binary image and we form a marker image F that is 0 everywhere, except at the image border, where it is set to 1-I that is:

$$F(x,y) = \begin{cases} 1\text{-}I(x,\ y) & \text{if } (x,\ y) \text{ is on the border of I} \\ 0 & \text{Otherwise} \end{cases} \quad (1)$$

Then:

$$H = \left\lfloor R_{I^c}^{D}(F) \right\rfloor^{c} \quad (2)$$

is a binary image equal to I with all holes filled. This operation eliminates all the holes in the segmented image. The resultant hole filled image is shown in Fig. 2d.

Clumped cell detection using ellipse fitting: After post processing operations, the resultant image contains isolated cells and clumped cells representing the RBCs and WBCs. Since, most of the cells in the blood smear image have elliptical or circular shape, this study proposes an efficient methodology to differentiate the clumped cells from single cells based on an efficient ellipse fitting algorithm. The purpose of this algorithm is to model the contour of each connected component in the image with a suitable ellipse and compute the eccentricity of each fitted ellipse. Then, an optimum threshold based on the eccentricity of the ellipse is empirically decided to distinguish the connected components as clumped cells or single non-clumped cells.

The clumped cell detection algorithm contains two major steps. In the first step, for a given connected component $C_i = \{P_i^1, P_i^2, P_i^3, ..., P_i^n\}$ where $p_i = (x_i, y_i)$ is the contour point and n is the number of contour points an ellipse E_i is fitted on the contour points by an ellipse fitting algorithm based on direct least square method (Fitzgibbon *et al*., 1999). It is computationally efficient and provides robust results in the presence of noise and occlusions. After fitting the ellipse in the second step, we use the following criteria to decide the optimum threshold to detect the clumped cells in the image.

Given a certain connected component C_i of the image, let E_i is the ellipse fitted on that connected component. We denote the eccentricity of the ellipse E_i by E_c and is computed by finding the ratio of half the distance between the foci (c) to the length of the semi major axis (a). For any ellipse E_i where $0<E_i<1$ and we let :

$$\varphi = \begin{cases} 0 \text{ if } 0 \le E_c \le 0.5 \\ 1 \text{ if } 0.6 \le E_c \le 1 \end{cases}$$

where, E_c is the optimal threshold of the eccentricity of the ellipse E_i. C_i is considered as a clumped cell if φ is 1 and non-clumped cell otherwise. The algorithm detects all the clumped cells in the segmented image efficiently. The resultant images obtained after this procedure is shown in Fig. 2e and f.

RESULTS AND DISCUSSION

We have conducted the experiments on 200 malaria infected peripheral blood sample images collected from public health centers. These are RGB color images of 2048×1536 pixel resolution. Each image represents a section of microscopic field at 1000X magnification. The blood cell images are examined and annotated by specialists in malaria diagnosis by manually drawing the cell boundaries. These are used as ground truth for evaluation of the proposed algorithm.

Evaluation: The success of clumped blood cell detection algorithm mainly depends on how accurately the clumped cells are detected in the image which contains both clumped and non-clumped cells. In our experiments, we evaluate the results visually and quantitatively. The accuracy of the algorithm is evaluated by comparing our results with the ground truth of the images. For that we consider the ground truth of the clumped and non clumped cell boundaries obtained from the experts as gold standard and compare the boundaries of the detected clumped cells obtained from the proposed cell detection algorithm. Based on this comparison, we identified the number of proper detections and misdetections. The results for few images are given in Table 1.

Comparison: We compare our results against two methodologies which are commonly used for clumped object detection in cell image analysis. One is based on area based thresholding method where the areas of the connected components are computed and optimum threshold is decided to separate the clumped and non clumped cells. In the other method, the threshold is computed from the area of the convex hull of the connected components (Kong *et al*., 2011).

The method based on area thresholding did not provide desirable accuracy as the parasite affected enlarged RBCs and WBCs were falsely detected as

Table 1: Detection accuracy using ellipse fitting

Imges_name	No. of clump cells present	No .of clump cells detected	Ground truth by specialists	Accuracy of proposed method (%)
1	35	34	35	97.1
2	31	30	31	96.7
3	36	34	36	94.4
4	33	33	33	100.0
5	39	38	39	97.4

Table 2: Comparision of the proposed method with commonly used methods

Methods	Total cell clusters	Correctly detected (%)
Proposed ellipse fitting method	7550	95.8
Classical area thresholding method	7550	89.1
Convex hull based thresholding	7550	93.4

clumped cells as their size were comparable. In the convex hull based thresholding, the threshold value used was almost similar to the area based threshold to do the similar separation. While the proposed method provides considerable improvement over the area thresholding method and the convex hull based and can be observed in Table 2. In the proposed method, an optimum eccentricity threshold of 0.5 is used and we have obtained an improved accuracy which was primarily due to the right interpretation of the shape of the clumped cell and excluding enlarged RBCs /WBCs which are tending to be of circular shape. The comparison results are given in Table 2. We can observe from the result that the proposed method performs significantly better than the other two algorithms with considerable reduction in the misdetections.

CONCLUSION

In this research, we have presented an efficient clumped cell detection algorithm which incorporates effective image segmentation and two post-processing techniques useful for thin peripheral blood smear images which contain large number of clumped blood cells. Even though the aim of this entire research is to develop an automated malaria diagnosis system, the proposed clump cell detection algorithm is of great importance not only in the automation of malaria but also all studies related to cell image analysis. As seen in the results, it has been observed that the proposed algorithm can successfully detect the clumped cells with considerable increase in accuracy. The performance of the proposed method is evaluated and compared with other clump cell detection algorithms. The comparison shows that the efficiency of the proposed method is significantly higher than the other algorithms which are commonly used for detection of clumped cells.

RECOMMENDATIONS

The future aim of this research is to evaluate the performance of the proposed methodology using a large dataset which contains images collected from different public health care centers and images acquired using different imaging equipments. Our endeavor is to apply this algorithm to study the cell behaviors related to various hematological and histopathological applications as a part of our future research.

REFERENCES

Chan, T.F. and L.A. Vese, 2001. Active contours without edges. IEEE Trans. Image Process., 10: 266-277.

Cheng, J. and J.C. Rajapakse, 2009. Segmentation of clustered nuclei with shape markers and marking function. IEEE. Trans. Biomed. Eng., 56: 741-748.

Diaz, G., F. Gonzalez and E. Romero, 2007. Automatic clump splitting for cell quantification in microscopical images. Proceedings of the 12th Iberoamericann Congress on Progress in Pattern Recognition, Image Analysis and Applications, November 13-16, 2007, Springer, Valparaiso, Chile, pp: 763-772.

Diaz, G., F.A. Gonzalez and E. Romero, 2009. A semi-automatic method for quantification and classification of erythrocytes infected with malaria parasites in microscopic images. J. Biomed. Inf., 42: 296-307.

Fitzgibbon, A., M. Pilu and R.B. Fisher, 1999. Direct least square fitting of ellipses. IEEE Trans. Pattern Anal. Machine Intellig., 21: 476-480.

Kong, H., M. Gurcan and B.K. Belkacem, 2011. Partitioning histopathological images: An integrated framework for supervised color-texture segmentation and cell splitting. IEEE. Trans. Med. Imaging, 30: 1661-1677.

Kumar, S., S.H. Ong, S. Ranganath, T.C. Ong and F.T. Chew, 2006. A rule-based approach for robust clump splitting. Pattern Recognit., 39: 1088-1098.

Mao, K.Z., P. Zhao and P.H. Tan, 2006. Supervised learning-based cell image segmentation for p53 immunohistochemistry. IEEE Trans. Biomed. Eng., 53: 1153-1163.

Mumford, D. and J. Shah, 1989. Optimal approximations by piecewise smooth functions and associated variational problems. Commun. Pure Applied Math., 42: 577-685.

Ritter, N. and J. Cooper, 2007. Segmentation and border identification of cells in images of peripheral blood smear slides. Proceedings of the 13th Australasian Conference on Computer Science Vol. 62, January 30-February 2, 2007, Australian Computer Society, Inc, Ballarat, Victoria, Australia, ISBN:1-920-68243-0, pp: 161-169.

Ruberto, D.C., A. Dempster, S. Khan and B. Jarra, 2001. Morphological image processing for evaluating malaria disease. Proceedings of the 4th International Workshop on Visual Form, IWVF4, May 28-30, 2001, Springer, Capri, Italy, pp: 739-748.

Ruberto, D.C., A. Dempster, S. Khan and B. Jarra, 2002. Analysis of infected blood cell images using morphological operators. Image Vision Comput., 20: 133-146.

Tek, F., A. Dempster and I. Kale, 2005. Blood cell segmentation using minimum area watershed and circle radon transformations. Proceedings of the 7th International Symposium on Mathematical Morphology, April 18-20, 2005, Springer, Berlin, Germany, pp: 441-454.

WHO., 2017. Malaria: WHO child health programme dries tears in the Democratic Republic of the Congo. World Health Organization, Geneva, Switzerland.http://www.who.int/malaria/areas/rapid_access_expansion_2015/features/access-health-care-drc/en/.

Wang, H., H. Zhang and N. Ray, 2012. Clump splitting via bottleneck detection and shape classification. Pattern Recognit., 45: 2780-2787.

World Health Organization, 1991. Basic Malaria Microscopy Part I-2. World Health Organization, Geneva, Switzerland..

Yang, X., H. Li and X. Zhou, 2006. Nuclei segmentation using marker-controlled watershed, tracking using mean-shift and Kalman filter in time-lapse microscopy. IEEE. Trans. Circuits Syst. I Regul. Pap., 53: 2405-2414.

Zhou, H. and K.Z. Mao, 2006. Adaptive successive erosion-based cell image segmentation for p53 immunohistochemistry in bladder inverted papilloma. Proceedings of the IEEE 27th Annual International Conference of the Engineering in Medicine and Biology Society EMBS05, January 17-18, 2006, IEEE, Shanghai, China, ISBN:0-7803-8741-4, pp: 6484-6487.