AN EMPIRICAL FRAMEWORK FOR AUTOMATIC RED BLOOD CELL MORPHOLOGY IDENTIFICATION AND COUNTING

Razali Tomari, Wan Nurshzwani Wan Zakaria and Rafidah Ngadengon
Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Malaysia,Parit Raja Batu Pahat Johor, Malaysia
E-Mail: mdrazali@uthm.edu.my

ABSTRACT

In blood tests analysis identification of Red Blood Cell (RBC) morphology and count the RBC number is crucial to diagnose any symptoms of blood related disease. In current practice, such procedure is executed manually by a pathologist under light microscope. As the samples increased, manual inspection become laborious to the pathologist and since visual inspection is subjective, it might lead to variation to the assessed samples. To overcome such a problem, an automatic method is proposed by utilizing image processing procedure. Initially RBC regions are extracted from the background by using a global threshold method applied on a green channel colour image. Next, noise and holes in the RBCs are abolished by utilizing a morphological filter and connected component labelling. Following that, geometrical information of the RBCs' area is extracted to determine single and overlapping RBC region. The former region is further process to identify its morphology either normal or abnormal by using geometrical properties and Artificial Neural Network (ANN), while the latter will undergo cell estimation stage by using Circle Hough Transform (CHT) to estimate the number of individual cells. The proposed method has been tested on blood cell images and demonstrates a reliable and effective system for classifying normal/abnormal RBC and counting the RBC number by considering an overlapping constraint.

Keywords: red blood cell, image processing, artificial neural network, circle hough transform.

INTRODUCTION

Human blood consists of three major components which are red blood cell (RBC), white blood cell (WBC) and platelets. Red blood cell is the major percentage in the human blood and responsible to various function such as transporting oxygen throughout the body, carrying carbon dioxide and waste products away from cells and tissue (Fox, 2009). Normal shape of the RBC is biconcave disks with about 7 µm in diameter and 2.2 µm thick.

From medical perspective, the RBC diagnosis contributes information about various blood related disease. For instance, the shape of RBC and its deformability has connection to the relevant disease such as Huntington's disease, Myalgic Encephalomyelitis (ME) and Multiple Sclerosis (MS) (Vromen, 2009, Wang, 2008, Wang, 2010). To ensure a proper and correct treatment can be given to the patient, an accurate diagnosis is very crucial. The diagnosis can be from counting the number of RBC until determining any abnormal morphology in the RBC. Lower RBC count indicates low of specific vitamin (Sharif, 2012, Zahir, 2006) and abnormal RBC morphology give a sign of anaemia, reduction of haemoglobin (a protein that bind with oxygen molecule in RBC), and secondary effect of several other disorder (Webster, 2004).

Normally, the blood sample is being processed in laboratory by using chemical electronic device called haemocytometer or haematology analyser. Such procedure is very dependent on the lab technologist's skill by viewing the sample through microscope. Despite its long clinical success, this method requires an expertise to manually count and classify the cells which is tedious, time-consuming and qualitative process (Venkatalakshmi, 2013). In addition, the existing method contributes to inaccuracy, inconsistency and poor reliability diagnosis that may lead to false diagnosis situation. In order to overcome the problem, an image processing technique is increasingly recognized as a very useful technique for the automated RBC diagnosis procedure. The method consist of standard framework which are image segmentation, image enhancement, RBC counting by considering clumped area and RBC classification (Habibzadeh, 2013, Mohammed, 2014, Tomari , 2014, Tomari 2015).

Image segmentation and enhancement is a crucial step to prune out the RBC shape accurately to ensure smoothness in the higher level processing. It is a difficult and challenging problem due to the complex nature of the cells and uncertainty in microscopic image (Kim, 2001). For example the region between RBC, WBC and background is sometime mixing together since both color components are very close to each other. Wrong clustering and scattering can lead to a similar color pixel between cell and plasma as a background. Consequently, contributing to unclear boundary between them (Rane, 2014, Chinwaraphat, 2008). In order to diminish noise as a result of segmentation, a morphological operation such as dilation and erosion were widely applied. Dilation adds pixels to the boundary of object in the image while erosion removes the pixels of boundary (Hamghalam, 2009, Angulo, 2003). Morphological operators also include a

few steps, which are filling holes, area calculation, template calculation, opening, closing, and reconstruction. Mathematical morphological operators used to segment RBC by eliminating WBC appearance (Ruberto, 2000, Adagale, 2013). It is also used for extracting image components and useful information in representing or describing the region of shape such as boundary, skeleton and texture.

The RBC classification can be generally formulated as pattern recognition problem. Several studies have reported that Neural Network method as an effective classifier than statistical pattern classification methods (Kim, 2001, Jambhekar, 2011, Veluchamy, 2012). A variety of different neural network techniques are used to classify RBC. Its good reputation makes the Neural Network as one of the widely used classifier for RBC classification systems. Counting the RBC by considering the clumped region is very challenging problem. Several methods have been proposed to solve such an issue. Some of them are based on Watershed and Distance transform (Huang, 2010, Sharif, 2012), morphological (Angulo, 2003), Appearance model (Rongtai, 2012), distance information (Hamghalam, 2009) and Hough transform (Venkatalakshmi, 2013). Hough transform has shown more prominent performance for predicting number of overlapping cell as compared to distance transform (Tomari, 2015).

In this paper discuss an empirical framework to automate the detecting of normal and abnormal RBC, and eventually count the number of each cluster by considering an overlapping constraint. The paper is organized as follows: Section 2 presents the architecture of the system; Next, Section 3 shows experimental results with discussion; and finally, the conclusions and future research are presented in Section 4.

SYSTEM OVERVIEW

The method pipeline for automatically classify and count the number of RBC is outlined in Figure-1.
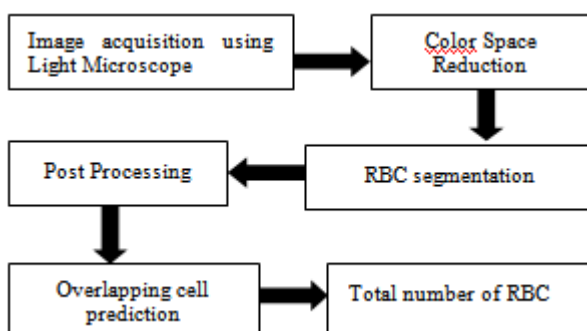

Figure-1. Method pipeline.

It is operated by first acquired input image from a light microscope that attached with an eye piece static

camera. Initially, the captured RGB image is converted into a single component color representation to make it convenient for the next processing. Next, the foreground is distinguished from the background by using adaptive global threshold method follow by low-level image post-processing methods in order to create a solid and noise-free foreground pixel map. Later, the connected regions of the foreground map are grouped together to identify overlapping and non-overlapping region. For the former region, further analysis is performed to predict the number of connected cells by using Hough transform and for the later the cell is classify between normal and abnormal by using combination of geometrical properties and neural network . Towards the end, performances of the method are evaluated based on its accuracy to classify the RBC and counting the number of each cluster correctly.

Color space reduction

Appearance of the RBC with respect to its background plays an important role. The more significant the appearance, the better it will be and generally RGB color information only gives just small effect. Because of that, the RGB image is transformed into a single channel color representation for an efficient computation in the higher level step.
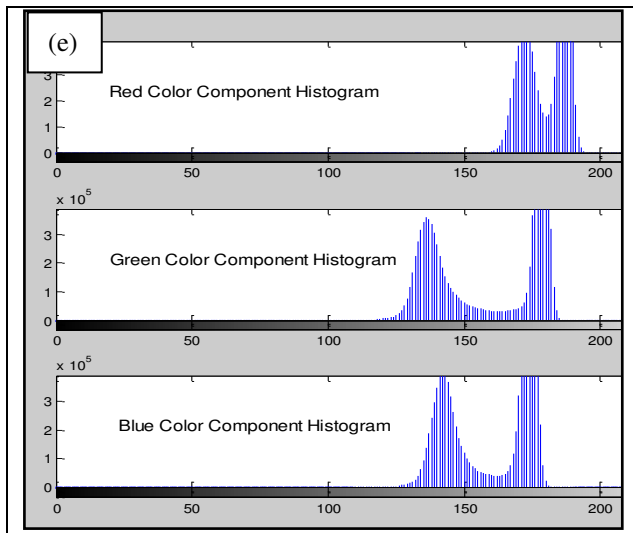
www.arpnjournals.com
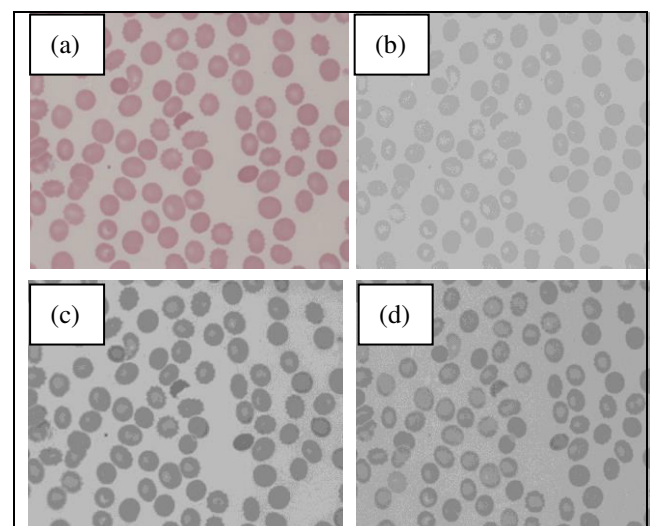
Figure-2. Color component selection. (a) RGB image (b) Red color component (c) Green color component (d) Blue color component (e) Color components histogram

In this project, the individual component of red, green and blue channels are investigated to determine an optimal color channel that capable to distinguish between red blood cell and the background. Figure 2 shows sample of the results where (a) is the RGB image, (b),(c) and (d) indicate the corresponding component of red, green and blue in the RGB image and (e) illustrate the histogram of the corresponding color components . It can be seen that, the red color component unable to precisely distinguish between the RBC and the background while the green and channel able to produce a better result. However, among the green and blue channel, qualitatively the green component gives the best contrast (wide histogram distribution) between the RBC and the background, and hence was selected for segmenting the RBC in the next section.

RBC segmentation and post processing

The image segmentation process is mainly implemented to partition an image into a region of homogenous representation corresponds to the object of interest in the image. In this paper, an Otsu adaptive threshold strategy (Otsu, 1979) is applied in the green channel of the RBC image to separating between two classes of region. This method works by finding threshold value that minimizes the weighted within class variance. Sample of the produced binary image as a result of Otsu segmentation process of green channel image (Figure-3 (a)) is shown in Figure-3 (b). It can be seen that, even though the attained image capable to detect the RBC region, there are still some noise and holes exist. To overcome such problems, a series of post processing method is applied.

The post processing aim is to remove noise and unwanted foreground cell from the segmented image. This process is crucial since the noise can significantly affect system ability to determine the RBC region accurately. In this paper, we use three methods which are morphological operation, connected component labelling (CCL) and bounding box filter to remove the unnecessary items.

Morphological operation works on binary image to change the size, shape, structure and connectivity of objects by using a structuring element and a set operator define by Erosion and Dilation. Erosion plays the role to 'shrinks' and 'thins' objects in image while dilation used to 'grows' and 'thickens' objects in image. The combination of both operators can be used to removes, break connection, clearing border and filling up holes. In this project, a sequence of two times of Erosion, two times of Dilation and contour filling algorithm is used to diminish the small noise and holes inside the cell. Sample output as a result of such process is shown in Figure-3 (c). It can be seen that, solid cell's shape is attained and small noise was successfully eliminated. Once such cells are in hand, the object candidates are labeled via connected component labeling (CCL) as can be seen in Figure-3 (d). The bounding boxes indicate the minimum and maximum rectangular cell location in the image. Since cell object in the border does not provide valuable information, it was removed by detecting minimum and maximum x and y bounding box locations that touch the image boundary (Red rectangle in Figure-3 (d)). Figure-3 (e) shows result after performing the mentioned process where the cell in border was successfully removed. Finally, the single and clump cells are identified by calculating the cell area. The former cell can be classified and count directly while the latter will be processed further to estimate the number of cells in the clumped area.
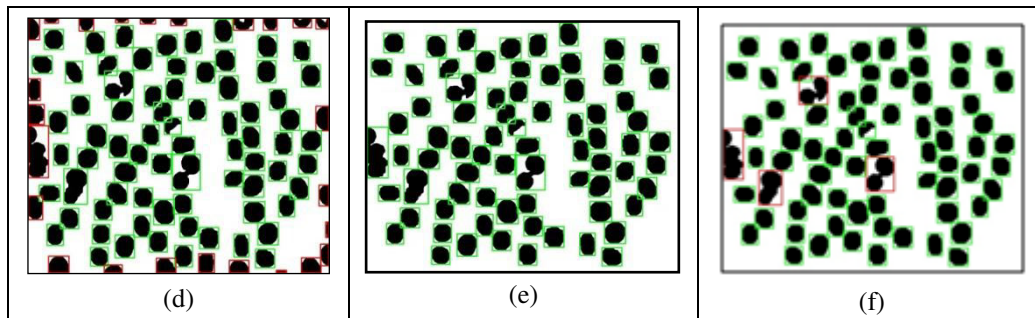

(a)   (b)   (c)

www.arpnjournals.com
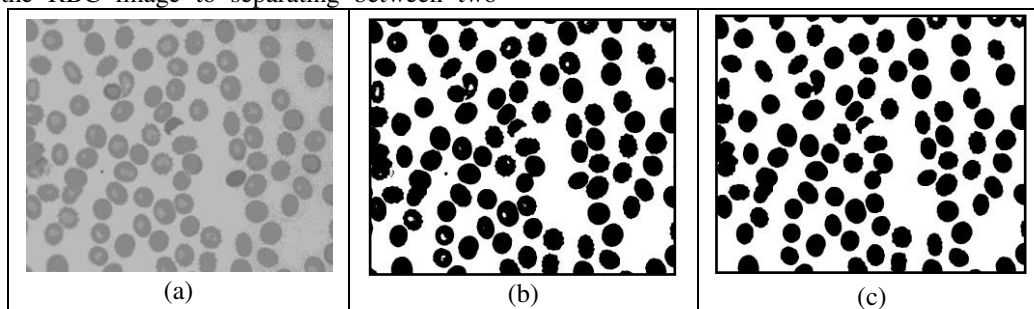
Figure-3. (a) Green channel image (b) Segmentation using Otsu Method (c) Morphological operation (erosion dilation and region filling) (d) Cell in border identification (e) cell in border removed (f) Detected overlapping cell in red bounding box.

Normal and abnormal single RBC classification

For classifying the single RBC, geometrical information which are compactness and moment invariant are extracted from the samples of normal and abnormal cell as shown in Figure-4. Compactness is a common shape measurement technique based on the ratio of object's perimeter object's area. The compactness provides information about how the object is form in term of the smoothness circle shape. When the compact value become higher, it shows that the objects shapes is more complex. Such property is really useful to distinguish between the cell types since for our case the abnormal cell shape is quite complex compare to the normal cell. However relying on a compactness value alone is not enough since when cell shape becomes more oval, the compactness value become higher even though the shape is not complex. Because of that, a second feature information is extracted which are moment invariant values.

Moment method has been used in the analysis and recognition of object shape. One advantage suggested by moment method is that it can easily equipped to be invariant in 2D transformation such as translation, rotation, reflection and scaling. This property is very convenient since usually the RBCs are distributed in random position and orientation. Furthermore, because the invariant do not provide any or less information than the original moment values, the features required for classification is reduced, and thus decrease the complexity of the learning problem. For this project, we use seven HU moment features (Hu, 1962) to represent the RBC shape.
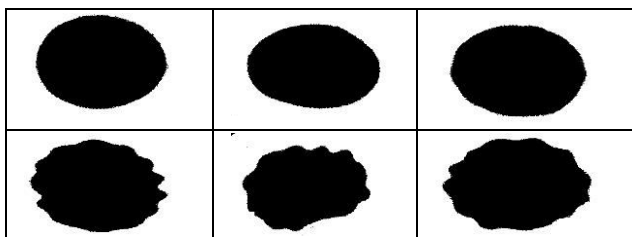

Figure-4. Sample of normal (top) and abnormal RBC (bottom).

Once the objects information is extracted, the next stage is to classify between the RBC classes. To satisfy such requirement, the robust Artificial Neural Network (ANN) classifier is selected. The ANN is a mathematical approximation of biological brain and has been identified as a useful framework for precise modelling of nonlinear response. It comprises a number of neurons connected together to form a network. The weights that linked between the neurons are where the functionality of the network resides. Before the network can be useful it needs to be trained. The training session will find an optimal weights value that minimize error between the inputs and the targets. In this paper, Levenberg Marquardt algorithm with mean square error (MSE) cost function is used as the learning procedure. Data from the normal and abnormal RBC features which are compactness and HU moments invariants are fed to the ANN module during the training process. The networks setting is consider optimal when achieved highest recognition rate in both training and validation data set.

Cell estimation from clumped RBC area

Splitting the clumped RBC is important in order to gain any related information in the 'covered' area. In this paper Hough transform is used to satisfy such requirement. Hough transform (Duda, 1972) is a feature extraction technique used in computer vision for finding lines, circles or other simple form in an image. It was initially suggested as a method for line detection in edge maps of images and then extended to detect the general low-parametric objects such as circles. To estimate the RBC in the clumped region, only an estimation of circle is required since basically the cell shape is lies in between circle shape to oval shape. For that reason, this paper focusing only on Hough Circle Transform (HCT).

HCT works by initially converting the input image into edges via canny edge detector (Canny, 1986). Following that, local gradient is calculated in every nonzero point in the edge maps by using Sobel filter. Using the gradient, every point along the line indicated by this slope, from a minimum to maximum is incremented in the accumulator. The candidate center is then selected

from the accumulator points that provide the highest value. The region of clumped cell image is fed to the HCT algorithm for cell estimation procedure. To determine an optimal circle number, there are four parameters that need to be tune which are accumulator threshold value, minimum distance that must exist between two circles to consider them as distinct circle, minimum radius and maximum radius. In our work, such parameters are empirically determined to be 19, 12, 8 and 29 respectively. Sample of the estimated cell is shown in Figure-5. It shows that, the HCT capable to estimate the number of circle in the clumped region that constitute to the single RBC cells.
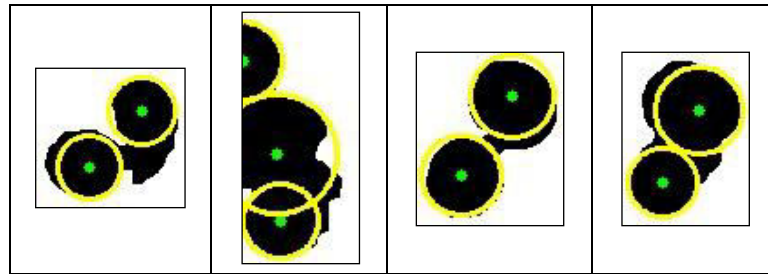

Figure-5. Hough circle estimation of the clumped cells.

EXPERIMENTAL RESULT AND DISCUSSIONS

In this section, performance of the RBC classification and counting system is assessed using four samples of smeared blood images labelled as im_1, im_2, im_3, and im_4 as shown in Figure-6. The evaluation framework consists of two different parts which are classifier performance and overlapping cell estimator performance.
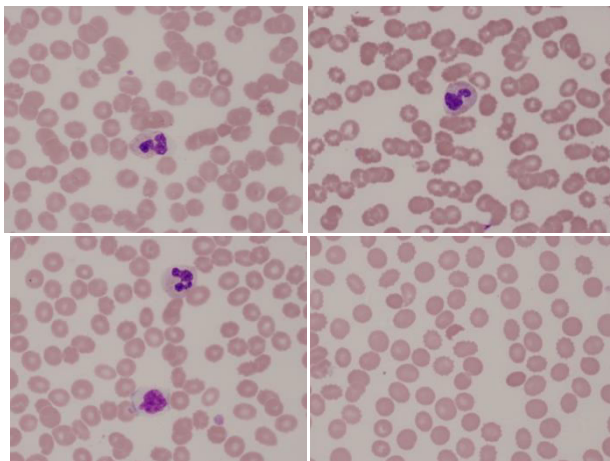

Figure-6. Images used for evaluating system performance. From top left to bottom right (im_1, im_2, im_3, im_4).

In the first part of assessment, we present an experiment conducted for assessing the RBC classification performance. Neural Network model from (Tomari, 2014) which consists of two input nodes, four hidden nodes with sigmoid activation function and one output node with sigmoid activation function is used in the experiment. Classifier performance is assed based on its ability to correctly identify the number of normal and abnormal RBC. For each class the quantitative measurement is performed based on True Positive (TP), False Positive (FP), True Negative (TN), and False Negative parameters. Eventually the precision (PC), recall (RC), and accuracy (AC) of the system are calculated by using equation (1). Precision provides information about how many of the detected fraction cells are correct, and recall tells how many cells are correctly detected in each class from the whole image. The accuracy on the other hand, gives evaluation about how well the overall system performance with respect to the ground truth data.

$$\mathrm{Pr}ecision = \frac{TP}{TP+FP} \; ; \mathrm{Re}call = \frac{TP}{TP+FN} \; ;$$

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN}$$

(1)

Table-1 summarizes our results. Overall the proposed method performs quite well with an average accuracy of 83%, 82% of average normal precision and 76% of average recall. This means that most of the object classes are correctly identified with an acceptable error rate. It also can be seen that the system provide a good detection result for the abnormal RBC detection compare to the normal one. This is due to the fact that the complexity of the abnormal RBC boundary is quite significant and easy to distinguish. On the other hand, the normal RBC is usually misidentified because of the imperfection in segmentation and post processing process. In the meantime, images im_1 and im_3 gave the lowest performance in term of recall rate with 50% and 23% respectively. The main reason for the lower recall rate is that the acquired images are quite blurred and hence it is quite confusing to distinguish between both cells.

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

Table-1. Result for identifying and count the number of normal (NM) and abnormal (ABN) RBC.

	RBC Type	TP	FP	TN	FN	PC (%)	RC (%)	AC (%)
im_1	NM	2	1	15	2	67	50	85
	ABN	15	2	2	1	88	94	
im_2	NM	10	1	19	5	91	67	83
	ABN	19	5	10	1	79	95	
im_3	NM	3	1	26	10	75	23	73
	ABN	26	10	3	1	72	96	
im_4	NM	25	3	39	2	89	93	93
	ABN	39	2	25	3	95	93	

In the second part of the assessment, an experiment is performed to evaluate the HCT performance to correctly estimate the number of cell including the one in the clumped region. The obtained value is compared with respect to the manual counting value. Table-2 summarizes the results. It can be seen that HCT capable to correctly count the number of cell with 94 % of average performance. This means that HCT shows a good performance for estimating the number of cells in the clumped region with an acceptable false alarm error. HCT tend to fail only when the RBC shape looks like an ellipse and when the cells overlapped too close with each other.

Table-2. Result of total RBC counting by estimating the clumped regions using Hough transform.

Image	Hough Transform	Manual Count	Accuracy (%)
im_1	94	110	86
im_2	93	103	90
im_3	85	87	98
im_4	78	78	100

Figure-7 shows example of output image result of the RBC classification and counting by considering an overlapping constraint. The single cell is denoted by a blue boundary while the clumped cell is mark by red boundary. For the single cell, the abnormal RBC was labeled by a black font while the normal RBC was labelled by a blue font. Meanwhile, for the clumped region the yellow circle indicate the estimated number of cell in such region by using Hough transform. The results show a promising outcome of the proposed system for completing its task in the captured images.
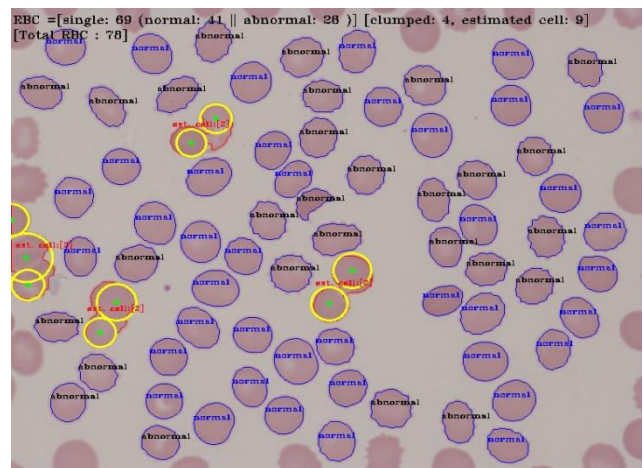
Figure-7. Sample result of cell classification and counting by considering an overlapping constraint of im_4.

CONCLUSIONS

In this paper, an empirical framework for automatically classify the RBC morphology and count the number for each cluster by considering the clumped region is proposed. The system consist combination of four main blocks which are segmentation and processing block, feature extraction block, classification block and clumped region estimator block. Each of algorithms in blocks gave a good performance during task completion with an acceptable error. We have shown that the ANN frameworks with four hidden nodes equipped with input from compactness and seven HU moment invariant can effectively identified the normal and abnormal cells with an average of 83% accuracy, 82% average precision and 76% of average recall. Meanwhile, counting the cell number by considering the clumped regions, the Hough transform shows a good performance with an average accuracy of 94%.

In future, the system can be improved by analyzing more sample cell to determine general

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

feasibility of the framework and including a method to identify the white blood cell region.

REFERENCES

Adagale, S. S. and Pawar, S. S.2013. Image segmentation using PCNN and template matching for blood cell counting. IEEE Intl. Conference on Computational Intelligence and Computing Research (ICCIC), pp. 1-5.

Angulo, J., and Flandrin, G. 2003. Automated Detection of Working Area of Peripheral Blood Smears using Mathematical Morphology. Analytical cellular pathology, 25, pp. 37-49.

Canny, J., A. 1986. Computational Approach To Edge Detection. IEEE Trans. Pattern Analysis and Machine Intelligence, 8(6), pp. 679-698.

Chinwaraphat, S., Sanpanich, A., Pintavirooj, C., Sangworasil, M., and Tosranon, P. 2008. A Modified Fuzzy Clustering for White Blood Cell Segmentation. 3rd International Symposium on Biomedical Engineering, pp. 356-359.

Duda, R. O. and Hart, P. E. 1972. Use of the Hough Transformation to Detect Lines and Curves in Pictures. Comm. ACM, 15, pp. 11-15.

Fox, S. I. 2009. Human Physiology. Mc Graw Hill International Ed., 11th Ed., New York.

Habibzadeh, M., Krzyżak, A., and Fevens, T. 2013. Comparative study of shape, intensity and texture features and support vector machine for white blood cell classification. Journal of Theoretical and Applied Computer Science, 7(1): p. 20-35.

Hamghalam, M., and Ayatollahi, A. 2009. Automatic counting of leukocytes in giemsa-stained images of peripheral blood smear. International Conference on Digital Image Processing, pp. 13-16.

Huang, J. 2010. An Improved Algorithm of Overlapping Cell Division. International Conference of Digital Object Identifier. Intelligent Computing and Integrated System, pp. 687-691.

Hu, M.K. 1962. Visual Pattern Recognition by Moments Invariants. IRE Trans. Info. Theory, vol. IT-8, pp. 179-187.

Jambhekar, N. D. 2011. Red Blood Cells Classification using Image Processing. Science Research Reporter, 1, pp. 151-154.

Kim, K., et al. 2001. Automatic Cell Classification in Human's Peripheral Blood Images Based on Morphological Image Processing. Advances in Artificial Intelligence, pp. 225-236.

Mohammed, E. A., Mohamed, M. M., Far, B. H., and Naugler, C. 2014. Peripheral blood smear image analysis: A comprehensive review. Journal of pathology informatics, 5.

Natsution A. M. T, and Suryaningtyas, E. K. 2008. Comparison of Red Blood Cells Counting using two Algorithms: Connected Component Labelling and Back projection of Artificial Neural Network. IEEE Phhotonics Global@ Singapore, pp. 1-4.

Otsu, N. 1979. A Threshold Selection Method from Grat-Level Histograms. IEEE Transaction on System, Man and Cybernetics. 9(1), pp. 62-66.

Pedro, F. F. and Daniel, P. H. 2012. Distance Transforms of Sampled Functions. Theory of Computing, 8, pp. 415-428.

Poon, S. S, Ward R. K., and Palcic B. 1992. Automated Image Detection and Segmentation in Blood Smears. Cytometry, 13, pp. 766-774.

Rane ,K., P., Zope, G., D. and Rane, A. 2014. Blood Cell Segmentation-A Review. International Journal of Advanced Electronics and Communication Systems.

Rongtai, C., QingXiang, WRui, Z., Lijuan, F., and Chengmei, R. 2012. Red blood cell segmentation using Active Appearance Model. IEEE 11th International Conference on Signal Processing (ICSP), pp. 1641-1644.

Ruberto, C. Di, Dempster, A., Khan, S. and Jarra, B. 2000. 15th International Conference on Pattern Recognition, pp. 397-400.

Sharif, J. M., Miswan, M., Ngadi, M., Salam, M. S. H., and Mahadi, M. 2012. Red blood cell segmentation using masking and watershed algorithm: A preliminary study. International Conference on Biomedical Engineering (ICoBE), pp. 258-262.

Tomari R., Zakria W. N. W., Jamil M. M. A., Nor F. M., Fuad N. F.N . 2014. Computer Aided System for Red Blood Cell Classification in Blood Smear Image. Journal of Procedia Computer Science, 42, pp. 206-213.

Tomari R., Zakaria W. N. W, Ngadengon R. and Wahab M.H.A. 2015. Red Blood Cell Counting Analysis by Considering an Overlapping Constraint, ARPN Journal of Engineering and Applied Sciences, 10(3), pp. 1413-1420.

Venkatalakshmi, B. and Thilagavathi. K. 2013. Automatic red blood cell counting using Hough transform. IEEE Conference on Information and Communication Technologies (ICT).

Vromen, J. and McCane, B. 2009. Red Blood Cell Segmentation from SEM Images. 24th International Conference Image and Vision Computing.

Veluchamy M., Perumal K., and Ponuchamy, T. 2012. Feature Extraction and Classification of Blood Cells using Artificial Neural Network. American Journal of Applied Sciences, 9, p. 615.

Wang, R., MacCane, B. and Fang, B. 2010. RBC Image Segmentation based on Shape Reconstruction and Multi-scale Surface Fitting. 3rd International Symposium of Information Science and Engineering, pp. 586-589.

Wang, R. and MacCane, B. 2008. Red Blood Cell Classification through Depth Map and Surface Feature. International Symposium on Computer Science and Computational Technology, pp. 339-342.

Webster, J. G. and Cazzanti, S. C. 2004. Bioinstrumentation-Hematology. John Wiley and Sons, Inc., 2004, ISBN978-0-471-26327-2, pp.170-188.

Zahir, S., Chowdhury, R. and Payne, G. W. 2006. Automated Assessment of Erythrocyte Disorder Using Artificial Neural Network. IEEE International Symposium on Signal Processing and Information Technology.