Red blood cells and white blood cells detection by image processing

Irwan Rahadi¹, Meechoke Choodoung², and Arunsri Choodoung³

¹Department of Statistics, Universitas Hamzanwadi, Lombok, Indonesia
²Department of Mathematics, Mahidol University, Bangkok, Thailand
³Center of Medical Laboratory Service, Mahidol University, Bangkok, Thailand
E-mail: irwanrahadi8@gmail.com

Abstract. The common method of red and white blood cells identification and counting consider the manual processes on microscope which is arranged by the laboratory’s technician with their own experience. In this research, we will develop a computer program to detect and identify the proposed objects based on their pattern. The proposed objects are Red Blood Cells (RBCs), and White Blood Cells (WBCs). For blood cells identification and classification, an idea of Viola and Jones will be followed. Adaboost (adaptive boosting) method will be applied to increase the accuracy of the error of learning algorithm. The output of the proposed program shows that all the types of cells mentioned can be detected and classify effectively by showing the number and time spent of cells detected.

1. Introduction
The development of technology helps humans to design ways to solve many problems related to health. One of them is using images as the main reference to detect disease. The images will be examined by a program developed to understand their characteristics. In this study, the author tries to conduct research in the field of white blood cells (WBCs) to detect cells [1].
In calculating the amount of blood in the human body, there are two ways to diagnose and filter, called the Complete Blood Count (CBC) and Differential Blood Count (DBC) [2]. CBC can be done automatically using a cytometer, while DBC requires further steps because the process is done manually by experts. This process allows the emergence of different results because it is done manually, mainly due to physical limitations, the level of concentration in question. Under these conditions, simpler steps are needed to help technical personnel identify and count cells. Image processing is the digital conversion of an image with several operations to draw some ideas or information about the image[3].
Study of image processing (IP) was started in the 1950s with a limited aim in cervical cancer mass screening and started in 1D specimen. One decade later, based on their calorimetric and morphological measurements, white blood cells (WBCs) was counted automatically using two-dimensional (2D) images. In the mid of 1970s, systems was developed for blood testing techniques which were supported by computer assistance. The development of Image processing research in 3D was established in the 1980s. Last ten years, image processing methods have been used as the principal technique numerous research in cell counting, cell types and shape identification[4][5].
Because human blood provides some information about their physical condition. This condition opens up more ideas for researchers to conduct research, especially in the effort to identify automatically which has an impact on the short time needed, the reduction in processing costs, results that can be more accurate and generally increases the attitude of anticipation for disease[6]. Blood cells are red colored...
 fluids that flow in the human body. These cells consist of white blood cells (WBC), red blood cells (RBC), platelets and plasma[1]. White blood cells or leukocytes are cells with nucleus and mitochondria [7] [8]. Leucocyte will be useful for the immune system. In addition, several types of cells that have similar functions in body protection are erythrocytes, neutrophils, monocytes, eosinophils, and lymphocytes [10]. Red blood cells or erythrocytes have the main function in the transport of oxygen and carbon dioxide and regulation of vasodilation [9].

This research focuses on identifying red blood cells (RBCs) and white blood cells (WBCs) based on their form. The introduction process is done by haar-cascade classifier by collecting data and adjusting the information that has been given. This research procedure will be explained as follows. The first part will provide information about related research articles that contain information about identifying white blood cells automatically and then counting the amount. Next will explain the research methodology, research results and discussion. In closing is a reference.

2. Research Methodology

In this part, we will provide some information which will be applied in the research.

2.1 Red Blood Cells

![Figure 1. Form of Red Blood Cells](image1)

![Figure 2. White blood cells among RBCs](image2)

RBCs is the highest number of blood composition, it covers around 40-45% of total volume of the blood [1][11] and plays the vital biological responsibility in oxygen and carbon dioxide circulation, besides in regulation and protection[7] [12]. RBCs become the most deformable cells in the human body with three features: geometry, cytoplasmic viscosity, and viscoelastic properties of RBCs membrane cortex structure [13]. The images on Figure 1 show the morphological form of RBCs.

On the study of red blood cell development, normal RBCs will be like donuts without a hole in the center, just only thicker edge and thinner in the middle. Since the various sizes of the vessel in our body cause these cells have the flexibility to change their shape to pass through the vessel. Generally, mature RBCs in the human body will stay in these conditions: diameter approximately 7-8µ with 80-100 femtoliters (fL) of internal volume, called "normocytic", the thickest point is about 2-2.5µ, 20-30 trillion erythrocytes in adult, lifespans about 120 days. RBCs contain hemoglobin, a red-orange appearance with a central pallor (no larger than 3µ), an iron-rich protein, with the ability to tie up Oxygen and deliver it to whole tissues and organs, this normal condition called normochromic [1] [10][14].

Oxygen will be useful in producing energy for surviving, and the residue, Carbon Dioxide, will be delivered back to lung and throne out during breath [1], while certain vitamins and minerals such as iron, copper, folic acid, and vitamin A, B6, B9, and B12 are necessarily RBCs formation. An unstable amount of them can affect the hemoglobin deficiency or RBCs reducing. Besides, provide information about Leukemia, the form of RBCs depicted some information whether our blood has enough oxygen based on hemoglobin proportion on cell images. Hemoglobin is a protein on RBCs composed by hemes which can bind the iron molecules to bind the oxygen and transport throughout the body. The reaction between iron and oxygen cause the red color on blood. We can draw some information from that's condition about the blood condition. In this situation, the researchers try to solve faster than manual
identification base on the form of hemoglobin shown on the cells. Therefore, image processing was involved to help the expert identify the cells faster and more accurate.

2.2 White Blood Cell
White blood cells (leukocytes) are cells that make up blood components. These cells function is to help the body against various infectious diseases as part of the immune system. White blood cells are colorless, have a nucleus, can move amoeboid, and can penetrate the capillary wall / diapedesis. Under normal condition there are 4x10^9 to 11x10^9 cells in a liter of healthy adult human blood - around 7000-25000 cells per drop. In each cubic millimeter of blood there are 6000 to 10000 (an average of 8000) WBCs. In the case of leukemia, the number can increase up to 50000 cells per drop. Leukocytes derived include: NK cells, prick cells, eosinophils, basophils, and phagocytes including macrophages, neutrophils, and dendritic cells. There are several types of white blood cells called granulocytes or polymorphonuclear cells, namely: Basophile; Eosinophil; Neutrophil, and two other types without granules in the cytoplasm: Lymphocyte; Monocyte.

2.3 Haar Cascade
This research will focus on RBCs and WBCs recognition and separation based on their characteristics. Therefore we will apply Haar Cascade Classifier which was started for face recognition and detection. Haar Cascade was the first time came from the idea of Viola Jones [15]. They have created the algorithms to detect the face. The algorithm inspired the researchers in other kinds of image identification.

2.3.1 Features. Haar Features or Haar-like feature has been popularly applied for object representation [16]. A Haar-like feature consists of two or more vertically or horizontally rectangular regions and its value is the difference between sums of pixels within these rectangular regions [18]. Viola and Jones created some features to analyze a given subwindows on images shown on Figure 3.

![Figure 3](image)

The application of Haar Features to detect the proposed objects The images above show how Haar feature will be applied to image identification, both to RBCs and WBCs. On this research, we differentiate the features base on their form

2.3.2 Integral Images. Integral Images also called Summed-Area Tables is a simple operation to calculate the sum of image's pixel by considering the rectangle features of that image. The calculation will be done by count the point of images intermediately.

![Figure 4](image)

The position of the grey rectangle is calculated by the operation of \((x_2,y_2)-(x_1,y_2)+x_2,y_1)\) + \((x_1,y_1)\)

The position of the grey rectangle is represented by a rectangular bounding box \(((x_1,y_1), (x_1,y_2), (x_2,y_1),(x_2,y_2))\) where \((x_1,y_1)\) is the start point on the top left corner. \((x_1,y_2)\) and \((x_2,y_1)\) are the points related to a width and a height of the rectangle respectively and \((x_2,y_2)\) is the new point based on the value of \((x_1,y_2)\) and \((x_2,y_1)\)[19].
2.3.3 AdaBoost. AdaBoost (Adaptive Boosting) is the boosting method to increase the accuracy of the error of learning algorithm [20]. The ability of this method is in error rate adaptability of the individual weak learners on each stage and using the average classification error to decide the final weighting. The key to the success of boosting is the method for incrementally selecting the weak learners and for re-weighting the training examples after each stage. The AdaBoost algorithm does this by re-weighting each sample as a function of whether it is correctly classified at each stage, and using the stage-wise average classification error to determine the final weightings among the weak classifiers. The AdaBoost method works as shown in Figure 5 [18].

2.3.4 Cascade. The result of Adaboost method is some stages with strong classifier[21]. Each stage work to identifies whether a given subwindow is classified as a non-red blood cell or red blood cell. If the subwindow is classified as the non-red blood cell, then it will be discarded automatically. Otherwise, the identified cell will be passed to the next stage of the cascade. This process will work on the same condition to detect WBCs. Cell identification process will work for all stage as long as the identified cells meet the condition of stages. it follows that the more maybe cell passed the stages, the higher the chance of the subwindow contains a cell. Each stage will provide strong classifier which is come from sample training and distinguishing positive and negative images with zero false. All strong classifier from stages will be applied as cascade classifier.

Figure 5. Adaboost method works

In cascade process, we adopt the guidance by Mahdi Rezaei[22]. Since we want to detect and separate RBCs and WBCs, then we apply two types of classifiers. The first one is to detect the RBCs, and the second one is to detect WBCs. The classifiers will work on same images. One classifier will detect the images contain RBCs, and other will detect the image contain WBCs. The result will be the number of each cell's category. To estimate the accuracy of the classifier, results will be compared to manual cell identification by the expert.

2.4 Counting of blood cells
Counting the number of RBCs and WBCs is done with the same ways. The algorithm was developed to detect the cells and count the number of them respectively. The detected cells are captured by the color...
square on the cell which is depicting the number of cells as well. Therefore, the algorithm should be
developed carefully and has high accuracy when applied.
To count the accuracy of our proposed method and manual count, we applied this equation:

\[
\text{Accuracy}_{\text{rbc}} = \left(1 - \frac{(\text{Cell}_{\text{rbc}} - \text{CellMC}_{\text{rbc}})}{\text{CellMC}_{\text{rbc}}} \right) \times 100\%
\]

(1)

\[
\text{Accuracy}_{\text{wbc}} = \left(1 - \frac{(\text{Cell}_{\text{wbc}} - \text{CellMC}_{\text{wbc}})}{\text{CellMC}_{\text{wbc}}} \right) \times 100\%
\]

(2)

Where \(\text{Cell}_{\text{rbc}}\) and \(\text{Cell}_{\text{wbc}}\) are representing the number of RBCs and WBCs cells detected by Haar
Cascade classification, \(\text{CellMC}_{\text{rbc}}\) and \(\text{CellMC}_{\text{wbc}}\) representing manual method.

2.5 Image dataset
The images of red blood cells will be treat to draw the best condition and grouped into two. One for
train data, and other for creating the algorithm. Image are proposed by Arunsri Choodoung, Center of
Medical Laboratory Service, Faculty of Medical Technology, Mahidol University. We have used the
dataset that includes two images in JPG dan BMP format with 24-bit color depth and resolution 600X.
Most of the images are captured by an optical laboratory microscope. The algorithm will be trained in
Python 3.5 and supported by OpenCV for image manipulation.

3. Research Results and Discussion
The given methodology is implemented in python 3.5 which was run in device with specification:
Intel(R) Core(TM) i3-2310M CPU @ 2.10GHz 2.10 GHz, RAM 8.00 GB, 64-bit operating system, x64
based processor. Implementation results are drawn from randomly selected images from the dataset. The
result of classification for RBCs and WBCs are discussed here. Further counting results are compared
to the manual counting of both categories to analyze the accuracy of the proposed method.

![Figure 6. (left) RBCs detected, (right) WBCs detected (blue square)](image)

3.1 RBCs and WBCs classification by haar cascade
RBCs and WBCs classification results by haar cascade are shown in the figures 5. The blue squares in
Figure 5(a) and (b) show the cells detected. Blue squares on (a) are representing RBCs detected, while
blue squares in (b) are representing WBCs detected. These forms can be counted automatically by
the command line in python programming.

3.2 Counting number of RBCs and WBCs
The result of proposed method are compared with a manual count for the same images. Based on the
equation 1 and 2, the number of normal RBC detected are 309 of 316 cells with accuracy is 97.8%.
WBCs detected are 1 of 1 cells with accuracy is 100%.

Conclusion
This paper provides a methodology to detect and count the normal and hypochromic red blood cells
automatically by Haar Cascade classifier and trained by the algorithm. The results indicated that the
algorithm can perform well as expected. The best detected can be done by adjusting the parameter on
the algorithm. Our proposed method is very effective in cell detection. Since we only need the red blood
cells images and train the classifier. The results will be drawn directly. It tends to work perfectly when
applied in RBCs detection center or the health center that need to detect the RBCs. Further research will be focused on the various types of blood cells. Such as red blood cell and white blood cells. This can be done by creating the classifier for each category to detect.

Acknowledgments
This research was performed under auspices of Department of Statistics, Faculty of Mathematics and Natural Science, Universitas Hamzanwadi. It was funded by Center for Education, Research, and Social Service, Universitas Hamzanwadi, Indonesia.

References
[1] Rohland L. Red blood cells (RBCs). Salem Press Encyclopedia of Health. 2017
[2] Ongun G, Halici U, Leblebicioglu K, Atalay V, Bekzac M, Bekzac S. An automated differential blood count system. InEngineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE 2001 (Vol. 3, pp. 2583-2586). IEEE.
[3] Dhingra MN, Kaur MH. Analysis of Leukemia Cell Extraction and Detection using Watershed. International Journal of Research Science and Management. 2014;1(7):6--10.
[4] Sayari G, Arpan R, Debasish S. Assessment of Morphologically Altered RBCs using Image Processing Tools. Materials Today: Proceedings. 2016; 3(10): 3361--3366.
[5] Sohrabinia M, Sadeghian S, Manavic D. Application of image processing and image analysis methods for large scale map revision. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2008;37:1351-1354.
[6] Sahastrabuddhe AP. Counting of RBC and WBC using image processing: A Hemoglobin. 2006; 12(15):14--17.
[7] Yeow N, Tabor RF, Garnier G. Atomic force microscopy: From red blood cells to immunoaematoology. Advances in colloid and interface science. 2017 May 11.
[8] Lias J. Analysis of red blood cell (RBC) classification using Ni vision builder Ai (Doctoral dissertation, Universiti Tun Hussein Onn Malaysia).
[9] Tomaiuolo G. Blood mimicking fluid for biotechnological applications. Fluidynamic behavior of red blood cells and droplets under confined shear flow (Doctoral dissertation, Universita degli Studi di Napoli Federico II). 2008.
[10] Gordon-Smith T. Structure and function of red and white blood cells. Medicine. 2013 Apr 1;41(4):193-199.
[11] Xie H, Zhang Y. The effect of red blood cells on blood heat transfer. International Journal of Heat and Mass Transfer. 2017 Oct 1;113:840-9.
[12] Blumers AL, Tang YH, Li Z, Li X, Karniadakis GE. GPU-accelerated red blood cells simulations with transport dissipative particle dynamics. Computer physics communications. 2017 Aug 1;217:171-179.
[13] Kim Y, Kim K, Park Y. Measurement techniques for red blood cell deformability: recent advances. InBlood cell-An overview of studies in hematology 2012. InTech.
[14] Al-Shura AN. Hematology in Traditional Chinese Medicine Cardiology. Elsevier; 2014 Apr 8.
[15] Wang YQ. An analysis of the Viola-Jones face detection algorithm. Image Processing On Line. 2014; 4: 128--148.
[16] Park KY, Hwang SY. An improved Haar-like feature for efficient object detection. Pattern Recognition Letters. 2014 Jun 1;42:148-153.
[17] Oren M, Papageorgiou C, Sinha P, Osuna E, Poggio T. Pedestrian detection using wavelet templates. In Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on 1997 Jun 17 (pp. 193-199). IEEE.
[18] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. InComputer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on 2001 (Vol. 1, pp. I-I). IEEE.
[19] Leang I, Herbin S, Girard B, Douzele J. On-line fusion of trackers for single-object tracking. Pattern Recognition. 2018 Feb 1;74:459-73.
[20] Freund Y, Schapire RE. A Short Introduction to Boosting*: *Journal of Japanese Society for Artificial Intelligence*. 1999; 14 (5): 771-780.
[21] Cen K. Study of Viola-Jones Real Time Face Detector.
[22] Rezaei M. Creating a Cascade of Haar-Like Classifiers: Step by Step. *Obtenido de University of Auckland*. 2014.