The morphological classification of normal and abnormal red blood cell using Self Organizing Map

R F Rahmat¹, F S Wulandari¹, S. Faza¹, M A Muchtar¹, I. Siregar²

¹ Department of Information Technology, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Medan, Indonesia

² Department of Industrial Engineering, Faculty of Engineering, Universitas Sumatera Utara, Medan, Indonesia

Email: romi.fadillah@usu.ac.id,

Abstract. Blood is an essential component of living creatures in the vascular space. For possible disease identification, it can be tested through a blood test, one of which can be seen from the form of red blood cells. The normal and abnormal morphology of the red blood cells of a patient is very helpful to doctors in detecting a disease. With the advancement of digital image processing technology can be used to identify normal and abnormal blood cells of a patient. This research used self-organizing map method to classify the normal and abnormal form of red blood cells in the digital image. The use of self-organizing map neural network method can be implemented to classify the normal and abnormal form of red blood cells in the input image with 93.78% accuracy testing.

1. Introduction

Along with the times, technology in the medical field is currently growing. Technology itself has become a necessity in all areas, especially biomedical. Many techniques have been developed that can help doctors in identifying a particular disease. Digital image processing can be implemented to detect uterine tumor or cancer, to determine diseases such as lungs disease, liver disease and bone disease, bone segmentation of other muscles, teeth classification and microscopic image analysis [1]. To identify a disease can be tested through a blood test, one of which can be seen from the form of red blood cells. The normal and abnormal morphology of the red blood cells of a patient is very helpful to doctors in detecting a disease [2].

Blood is an essential component of living creatures in the vascular space, due to its role as an intercellular communication media to various parts of the body with the outside world because of its function of carrying oxygen from the lungs to the tissues and carbon dioxide of the tissues to the lungs to be removed, bringing nutrients from the gastrointestinal tract to the tissue and then deliver the rest of the metabolism through organ secretions such as kidney, providing hormones from blood clotting materials [3].

Previous research that discuss the image of red blood cells are: determining the morphology of red blood cells using Artificial Neural Network (ANN) method [2], the calculation of overlap red blood cells using Morphology Operation method [4], calculating the number of red blood cells using Multilayer Perceptron [5], Detecting and calculating the amount of normal and abnormal form of red blood cells using the object feature extraction based on the shape [6].
Based on previous research we are encouraged to investigate the detection of normal and abnormal red blood cells by implementing the Self-Organizing Map method. In this study, we use Self-Organizing Map (SOM) or better known as Kohonen method to perform normal and abnormal red blood cells pattern recognition of an image. SOM is a method often used in clustering process. SOM included in competitive learning, where the output nodes will compete to be the winning node and become the only node to be activated. Also, the transformation can create group data automatically, similar data will be stored in the nearest cluster. Since the initial weights are determined randomly, the SOM can determine itself to enter a particular group. SOM also takes only a short time to complete the learning process. However, the success of the system using SOM depends on the size of the output value of neurons number, radius neighbors, and the pace of learning [7]. Related research on SOM or image classification includes: Classification of nutritional status of children [7], Vectors quantization on a 24-bit bitmap image compression [8], Rainfall prediction [9], character recognition [12], classification of ribbed smoke sheet using LVQ and shape feature extraction for its features [13] and character recognition using directional feature extraction [15].

We hope the system to be built will result in better accuracy and is able to help doctors to distinguish normal and abnormal red blood cells automatically. Then the analysis of normal and abnormal forms of red blood cells performed by a doctor is not always the same with other doctors. The accuracy and concentration of doctors significantly affect the analysis result, therefore required a system that can identify normal and abnormal forms of red blood cells with better accuracy and in a short time.

2. Research methodology

The general architecture of our proposed method is shown in figure 1.

![Figure 1. General Architecture](image_url)

2.1. Red Blood Cells

The most abundant cells in the blood membrane are red blood cells or also known as erythrocytes. Red blood cells in the form of biconcave discs with a diameter of about 7.5 microns, the thickness of the edges of 2 microns and the central part of 1 micron or less, are composed of very thin membranes that are very susceptible to diffusion of oxygen, carbon dioxide and cytoplasm, but do not have a cell
nucleus [10]. Viewed from the side, erythrocytes look like discs or biconcave with central acromial approximately $\frac{1}{3}$ - $\frac{1}{2}$ cell diameter [2]. Conventionally red blood cells are calculated by diluting a few drops of blood in a diluent solvent. The diluted blood is placed in the calculation space based on the known amount, and the number of cells presents calculated microscopically. Calculation of cells is assumed as the number of cells in one liter of blood. Calculation of red blood cells is considered as a number multiplied by $10^{12}$ per litter (i.e., $5 \times 10^{12}/l$) [10].

2.2. Data Pre-processing

2.2.1. Data Acquisition. In this research, we used several images which contain red blood cells inside. This image has been taken from Prodia Laboratory and several pathology doctors with their permission. The images taken from the laboratory has been classified into two general images. Images which contain abnormal red blood cells or we called it as anemia, and images which contain normal red blood cells. We collect 28 images that consist of 24 images which contain abnormal red blood cells and 4 images which contain normal red blood cells.

![2a](image1.png) ![2b](image2.png)

**Figure 2a.** Abnormal Red Blood Cells, **2b.** Normal Red Blood Cells

2.2.2. Greyscaling. It is the process of changing the pixel values from color (RGB) to gray-level or grayscale. The greyscale process is done by equalizing the pixel values of the three RGB values into one value [8][11][14]. The process of converting color image to grayscale image can be done using equation (1)

$$ I(x, y) = \frac{R + G + B}{3} $$

(1)

Where $I(x, y)$ is grey level in position $(x, y)$. While $R$, $G$, and $B$ declare the color component value of each color pixel value in position $(x, y)$ respectively.

2.2.3. Thresholding. Thresholding will generate binary image, an image that has two levels of gray which is black and white. In general, thresholding process of a grayscale image to generate binary image explained in equation (2):

$$ g(x, y) = \begin{cases} 1 & \text{if } f(x, y) \geq T \\ 0 & \text{if } f(x, y) < T \end{cases} $$

(2)

where $g(x, y)$ is binary image of grayscale image $f(x, y)$, and $T$ is the value of threshold. $T$ value holds significant role in thresholding process [6].

2.2.4. Erosion and Dilation. The erosion process is executed by comparing each pixel of input image to SE center value by superimposing the SE with the image so that the SE center is precisely at the position of the processed image pixel. Erosion process will generate smaller object. The hole in the
object will also appear to enlarge as it narrows the object boundary. The dilation process is executed by comparing each pixel of input image to SE center value by superimposing the SE with the image so that the SE center is precisely at the position of the processed image pixel.

2.2.5. **Boundary.** The boundary is the process to do segmentation on the images. We did background separation, and we fill the objects with different colors to differentiates them. In this process, we also do notation on the objects.

2.3. **Feature Extraction**

Invariant moment feature is beneficial for declaring objects by taking into account the object area. This feature uses central moment normalized basis. The moment generated can be implemented to handle translation, scaling and image rotation [13]. In this method, we took all 7 invariant moments as object features extracted from the image.

Another feature that we consider as a shape parameter is a roundness. Roundness describes the level of cellular roundness. Roundness is the ratio between object area and squared perimeter, which can be calculated using equation (13). The result is a value of ≤ 1. Value 1 indicates R object as a circle.

\[
R = (4\pi) \left( \frac{A(R)}{P^2(R)} \right) \tag{13}
\]

Where:

- R = Roundness
- A = Area
- P = Perimeter

2.4. **Self Organizing Maps (SOM)**

Self-organizing map (SOM) method or known as Kohonen, introduced by Teuvo Kohonen, is a competency-based neural network system capable of unsupervised learning due to its self-organizing ability. This neural network will study the distribution of set patterns without any previous class information [1].

3. **Result and discussion**

3.1. **Image Processing Results** – Here we describe the result from our proposed method based on our proposed general architecture in figure 1 above. Figure 3 shows that the process starts from original image, greyscale it down, then we did threshold to differentiate the background, then the morphological operator such as erosion and dilation are performed. The boundary process will be shown in figure 5. Figure 4 shows the extraction of the objects into 7 invariant moments and roundness values.
3.2. System Overview

The result of this research is in the form of image that displays the cells considered as normal and abnormal along with information on the number of red blood cells detected. From the identification results shown in Fig 5, the tested cell is considered to be normal that contains roundness value of 0.90. Red blood cells considered to be normal if the roundness value is approaching 1. From the sample shown in Fig 5 contained result of 7 invariant moment values:

1. 0.1602
2. 3.1938e-04
3. 1.2415e-06
4. 7.320e-09
5. 3.9522e-16
6. 7.4814e-11
7. -3.2081e-16

Figure 3. (from left up to right down). a. Original Image, b. Grayscaled Image, c. Threshold Image, d. Erosion Image, e. Dilation Image

Figure 4. Invariant Moment and Roundness Extraction
From the identification results shown in Fig 6, the abnormal cells contained roundness value of 0.86. The abnormal red blood cells tend to have a roundness value smaller than 0.9. The result depends on the input image.

From the sample shown in Fig 6 contained result of 7 invariant moment values:

1. 0.1689
2. 0.0024
3. 2.0364e-04
4. 4.9047e-06
5. -3.3754e-11
6. -1.0413e-07
7. 7.2269e-11
3.3. Testing Results

From the identification results shown in Fig 8, we present 5 samples from 15 testing images that we conducted before.

![Identification Results](image)

**Figure 8. Identification Results**

| Image Number | Cells in Total | Identified | Failed | Accuracy |
|--------------|----------------|------------|--------|----------|
| 1            | 23             | 22         | 1      | 95.65%   |
| 2            | 27             | 27         | 0      | 100%     |
| 3            | 28             | 26         | 2      | 92.85%   |
| 4            | 10             | 7          | 3      | 70%      |
| 5            | 16             | 10         | 6      | 62.5%    |
| 6            | 23             | 21         | 2      | 91.30%   |
We used 322 cell partitioned images from 15 big cell images. Based on our experiment shown in Table 1, it shows that the system successfully identified 302 cell images. From this result, we can accumulate overall accuracy from our system which is shown below.

\[
\text{Percentage of System Accuracy} = \frac{302}{322} \times 100\% = 93.78\%
\]

4. Conclusion
Based on the data testing done can be concluded that self-organizing map neural network method can be implemented to classify normal and abnormal forms of red blood cells with 93.78% accuracy. We also can conclude that the image pre-processing stage has not been able to separate the cells that coincide or overlap effectively. Therefore, we suggest several things for further research:
1. Self-Organizing Map (SOM) method can be compared with other methods in similar conducting research to get better results.
2. We built the system to identify normal and abnormal forms of red blood cells limited to the shape; we hope in future research, system will be developed to identify normal and abnormal forms of red blood cells using other two components which are based on size and color.
3. Several other methods are required in the pre-processing process to generate non-stacked blood cells. Thus, the next process can produce better data information.
4. There are still some cells considered as one cell, but in fact, there are two or more cells, for that issue, it will require a different method that can solve this problem.

5. References
[1] Putra D 2009 Digital Image Processing. Yogyakarta: Andi
[2] Warni E 2009 The morphological operation of red blood cells (erythrocytes) based on image processing and artificial neural networks Scientific Journal Electrical Engineering
[3] Tarwoto and Wartonah 2008 Medical Nursing Surgery of Hematology System Disorders. Jakarta: Trans Info Media.
[4] Usman K 2008 Calculation of Stacked Red Blood Cells Based on Digital Image Processing and Morphological Operation. National Seminar of Informatics, 31-38.
[5] Poomcokrak J and Neatpisarnvanit C 2008 Red Blood Cell Extraction And Counting. The Third International Symposium on Biomedical Engineering, 199-203
[6] Kartikasari D 2013 Abnormal and Normal Red Blood Cell Detection and Calculation System Based on Digital Image Processing Using Matlab. Thesis. Universitas Komputer Indonesia
[7] Zunairoh U, Khotimah B K, and Rachmad A 2012 Classification of Children Nutrition Status Using Kohonen Self Organizing Map. Journal of Bachelor of Informatics Engineering.
[8] Tae G F, Suwarno S, and Hapsari W 2010 Kohonen Self Organizing Map Implementation in Vectors Quantization on 24-bit Bitmap Image Compression. Jurnal Informatika, 6 pp 15-23.
[9] Septiadi D 2008 Soft Computing Application of Rainfall Prediction in Kalimantan. Journal
of Meteorology and Geophysics 9 pp 65 – 77

[10] Bain B J 2004 A Beginner's Guide to Blood Cells. 2nd Edition, pp. 9 - 22.

[11] Rahmat R F, Chairunnisa T, Gunawan D and Sitompul O S 2016 Skin color segmentation using multi-color space threshold International Conference on Computer and Information Sciences, ICCOINS 2016 – Proceedings pp. 391-396.

[12] Gunawan D, Arisandi D, Ginting F M, Rahmat R F and Amalia A 2017 Russian character recognition using Self-Organizing Map Journal of Physics: Conference Series 801, p. 12040

[13] Rahmat R F, Pulungan A F, Faza S and Budiarto R 2017 Image classification of ribbed smoked sheet using Learning Vector Quantization Journal of Physics: Conference Series 801, p. 12050.

[14] Nababan E B, Iqbal M and Rahmat R F 2016 Breast Cancer Identification on Digital Mammogram Using Evolving Connectionist Systems International Conference on Informatics and Computing (ICIC) pp. 132–136

[15] Seniman, Arisandi D, Rahmat R F, William and Nababan E B 2016 Chinese chess character recognition using direction feature extraction and backpropagation International Conference on Data and Software Engineering 2016