CHAPTER 3

Materials and Methods

3.1 Sample Preparation

The cells used in this study are peripheral blood mononuclear cells (PBMCs) which derived from the whole blood because percentage of CD4+ lymphocyte in PBMCs highly present. After isolating PBMCs from whole blood using Ficoll-Hypaque solution, 10% human AB serum is added to PBMCs for blocking FC receptors which highly express on lymphocyte's surface in order to prevent the non-specific binding. The incubation is done under 4°C for 30 minutes. After that, primary antibodies named MT4/3 MAbs (mouse anti-CD4) which bind specifically to the CD4 proteins on the T-helper cell surface and thus, no overlapped coupling with monocytes occur [31], are added in order to stain CD4 proteins. The reaction takesplace in 4°C for 30 minutes and then cells are washed by isotonic buffered diluent. Next, secondary antibodies named polyclonal rabbit anti-mouse immunoglobulins (anti-immunoglobulins Abs) tagged with PE are added in order to stain the CD4 protein-MT4/3 MAbs complexes and incubated for 30 minutes under 4°C in the dark to prevent the fluorescence fading. After performed washing, the FITC tagged anti-CD3 MAbs is added to stain CD3 proteins at 4°C for 30 minutes in the dark. Finally, cells are washed 3 times and preserved in 1% paraformaldehyde for maintaining the cross-linking between antigen and antibody. Briefly, CD4 proteins were stained by indirect staining, followed by CD3 proteins which were stained by direct staining technique. A schematic of cells stained by this procedure is illustrated in Fig. 3.1

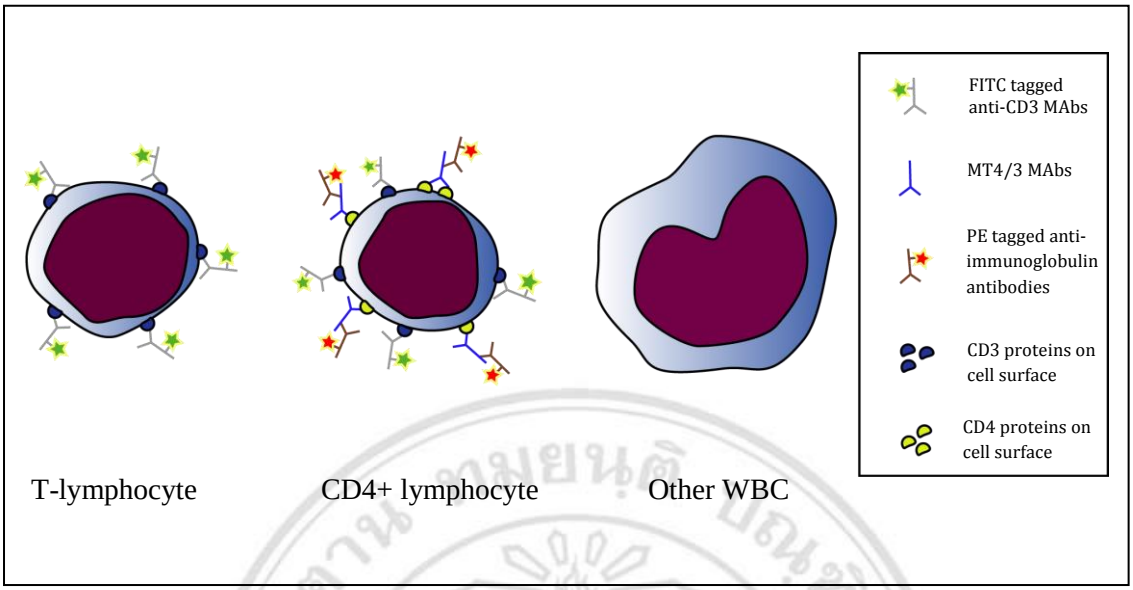


Figure 3.1 Schematic of stained cells.

3.2 Image Acquisition

PBMCs are visualized using an Olympus bx41 fluorescence microscope with 20x objective lens where the resolution is low. The purpose of using this magnification is to increase cell population while they can still be identified. Our images were captured under visible light, excitation/emission ~ 485nm/ ~519nm and excitation/emission ~490nm/ ~670nm for generating the bright field, green fluorescence and red fluorescence images respectively. One image set consists of three images captured with the same scene using an Olympus DP21 microscope digital camera contained 1/1.8 inches color CCD. The ISO is 200 with automatic exposure time and 0 exposure adjustment. The white balance is 55000K. Fig.3.2 demonstrates the cells which were stained by the procedure explained in 3.1 and visualized under bright field, green fluorescence and red fluorescence conditions. Among all cells, only CD4+ lymphocyte can be seen in three images.

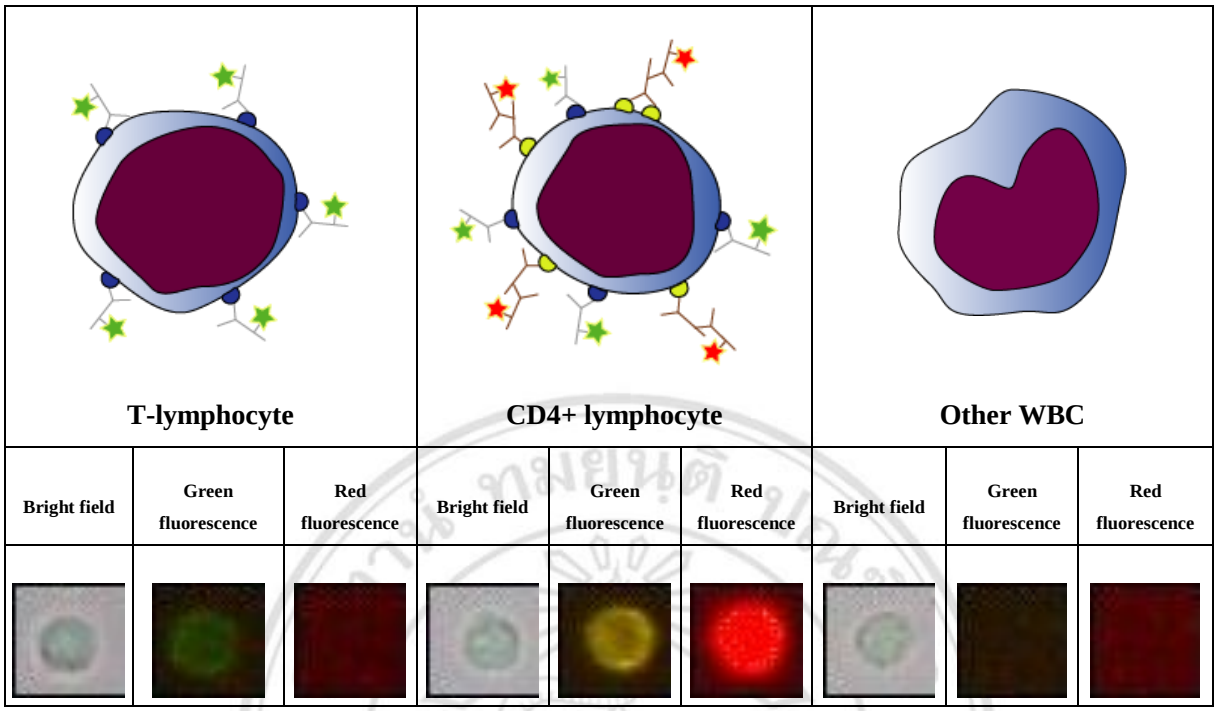


Figure 3.2 Stained cell imaging under fluorescence microscope.

3.3 Proposed Algorithm

According to the reviewing of previous study on cell detection methods, we are interested in developing our algorithm based on unsupervised approach. Although the supervised approach gave a very good result for cell detection, a training process is required. We want to avoid the complication of selecting appropriate samples for training system since our data set contains various objects, e.g. WBCs, debris, artifacts and red blood cells (RBC). So, in this study, supervised techniques are not taken into account.

As mentioned earlier, counting CD4+ lymphocytes requires 3 images; bright field, green fluorescence and red fluorescence images. Due to the different appearance of cells in bright field and fluorescence images, the algorithm is developed using the different method. Mainly, the proposed algorithm is divided into 3 contents. First, segmentation of WBCs in bright field image is done in order to find the possible location of CD4+ lymphocytes. Then, we detect color spots conveyed the interested cells in green and red fluorescence images. Finally, detected positions from bright field, green fluorescence and red fluorescence images are compared and thus, overlapping areas among three

images are defined as CD4+ lymphocytes. Diagram of our proposed algorithm is shown in Fig.3.3 and the detail is described below.

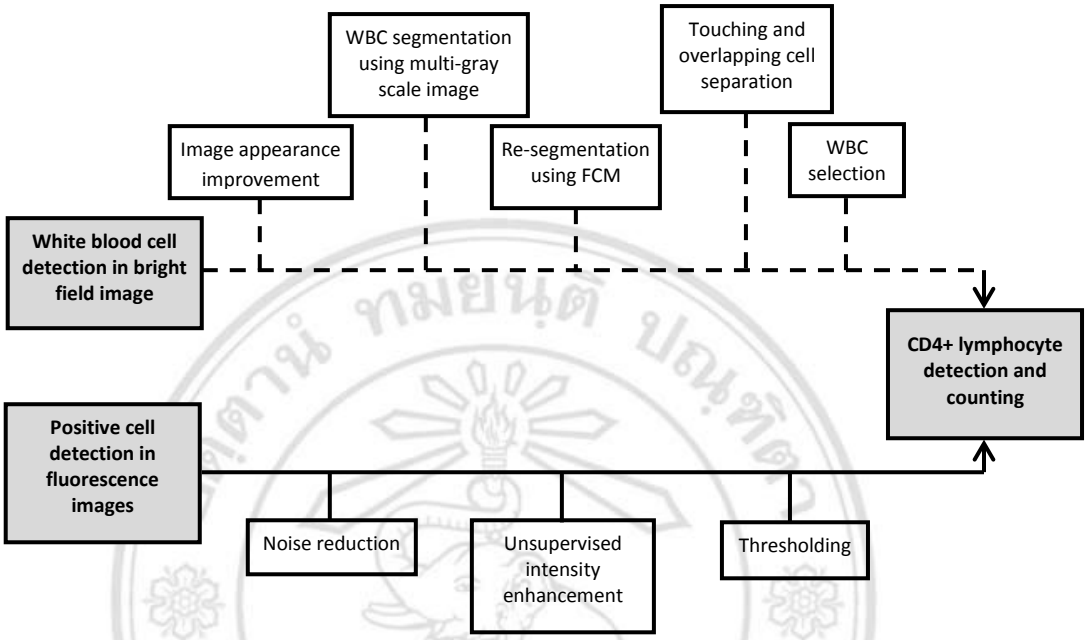


Figure 3.3 Diagram of proposed algorithm.

3.3.1 Algorithm for Detecting White Blood Cell in Bright Field Image

1) Image Appearance Improvement

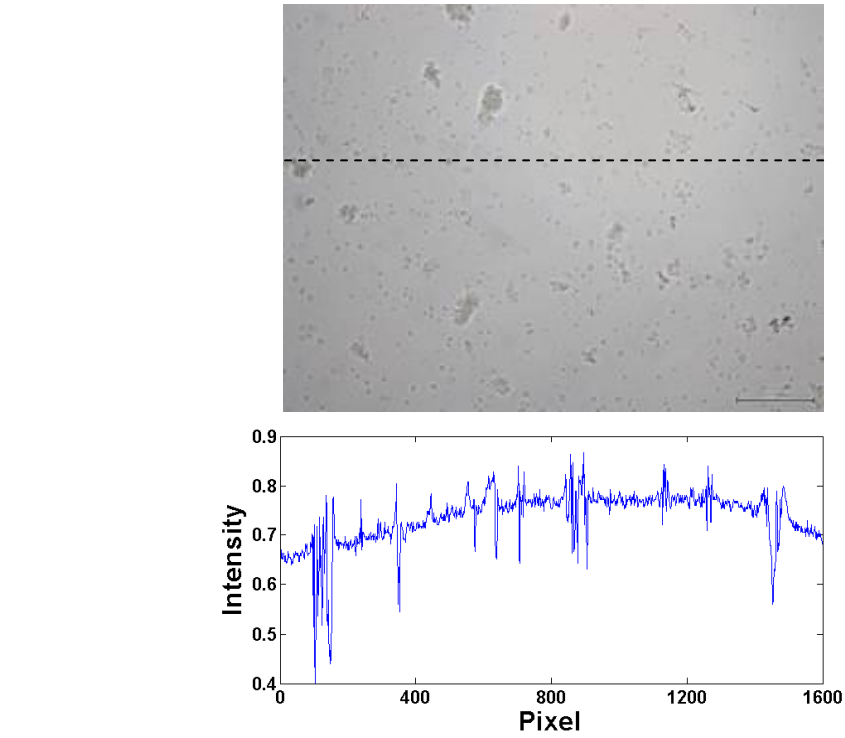
Since the image is captured under low magnification, the dispersed light from the light source apparently emerges. The 1-D plot of pixels along the single row (along the dash line) of the image to its intensity (Fig.3.4) demonstrates the brightness across the image. The curve plot (Fig 3.4a) corresponds to the characteristic of dispersed light which is lighter around the middle and gradually darker toward the image rim. Consequently, the intensity of the cells located around the center is higher than the outer. So, the appearance of the image should be improved to provide the appropriated input in order to increase the algorithm performance. There are many techniques for improving the image quality in spatial domain globally, such as contrast stretching [20] and locally, such as morphological contrast enhancement [10, 32] and adaptive contrast enhancement [25]. In our study, the global contrast enhancement is more suitable since the dispersed light effect is seen all over the image. The bivariate

Gaussian function is used to imply the characteristic of the dispersed light as demonstrated by the blue line in the Fig.3.4. Then the dispersed light effect can be eliminated by computing its invert version as denoted by

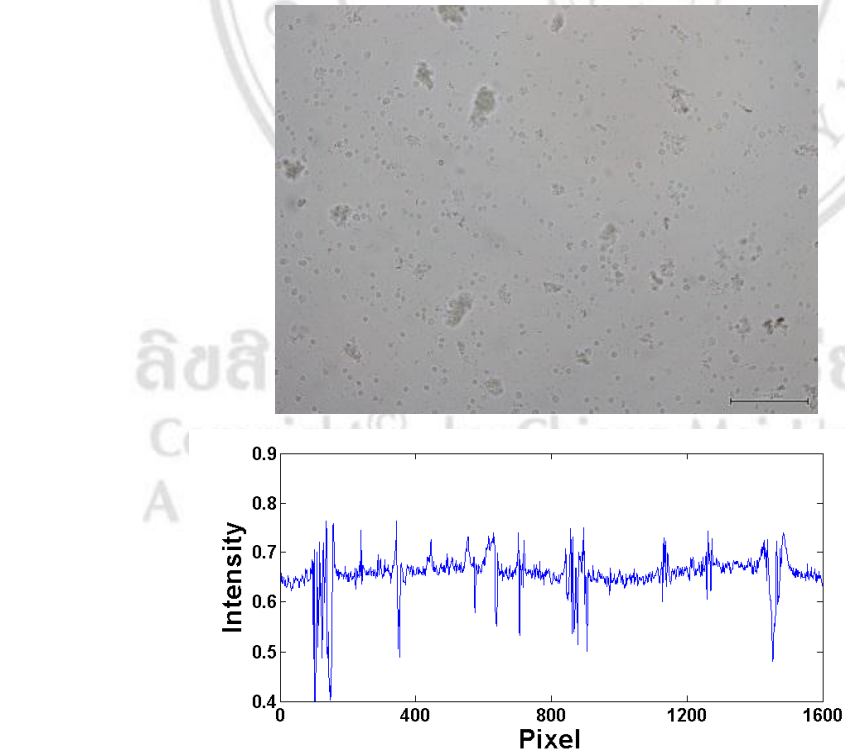
$$T(x,y) = 1 - Ae^{-\left(\frac{(x-x_0)^2}{2\sigma_x^2} + \frac{(y-y_0)^2}{2\sigma_y^2}\right)} \tag{17}$$

To define the parameters in eq.17, the gray scaled image $f(x,y)$ is first transformed into a binary image using the global thresholding method for highlighting the area Z where the dispersed light is located. The morphological operation was then applied to fit the contour shape. We assumed that the presented shape is an ellipse. Suppose that the ellipse has the coordinates $X = (x_1, x_2, ..., x_K)$ along x axis and $Y = (y_1, y_2, ..., y_L)$ along y axis which the center points x_0 and y_0 can be defined by the mean of *X* and *Y*, respectively. The variables σ_x and σ_y, which represent the spread of Gaussian function along the horizontal and vertical directions, are denoted by $\sigma_x = \max(X)$ and $\sigma_y = \max(Y)$. The amplitude A can be computed by the different of average intensity of the pixels in Z^C and *Z*. After substituting all variables in eq.17 in order to generate the inverse Gaussian function *T*, the output image *f '* is

$$f' = Tf \tag{18}$$



(a)



(b)

Figure 3.4 Intensity improvement. (a) Curve plot of intensities profile.
(b) Flatten plot after improvement.

2) WBCs Segmentation Using Multi-Gray Scale Image

Two groups of cells located on the different position in the same image are shown in Fig.3.5. Due to mildly different settlement of the cells in the specimen, even though they are in-focused, their appearance is different. The appearance of the cell boundary in Fig.3.5a (right) is worse than in Fig.3.5a (left) but the cellular internal color is greener and brighter. In this study, we use three types of gray scale images for cell representation, namely gray scale, R channel and chrominance images. We found out that the clearly-seen boundary is properly displayed by gray scale (Fig.3.5b (left)) and R channel image (Fig.3.5c (left)). Also a gradient caused by Halo ring is highly express. However, the ill-defined boundary cells cannot be displayed properly by the gray scale and R channel images since they show too bright content and poor boundary. Fortunately, due to the prominent bright green color, the contrast of the cell and background intensity is high when displayed by a chrominance image. Using three gray scale images can increase an opportunity to detect WBCs in bright field image. We applied the edge-based segmentation method to the gray scale and R channel images since the edge of the WBC is outstanding. On the other hand, intensity-based segmentation is applied to the chrominance image. Each gray image is operated separately at the step 1.2. Then, all resulting are merged at the end.

2.1) Edge-Based WBC Candidate Detection: Gray Scale and R Channel Images

Before applying edge detection, the image f 'should be blurred using median filter for noise cancellation. After that, the Canny method is used to detect edges in the image. After applying the canny method, a large amount of lines appear. We assume that clearly-seen boundary cells have closed edge which can be easily drawn out by performing a hole filling operation. However, some of them possibly have a broken and opened contour. So a dilation operator should be applied for gap closing before filling the holes again. The images obtained from the double hole-filling step are combined after discarding the objects whose size are out of the range 100-1000 pixels. This estimated range of values gave good results for our data. Also the objects that touch the border of the image are discarded. The remaining objects are segmented again using FCM clustering which will be explained later.

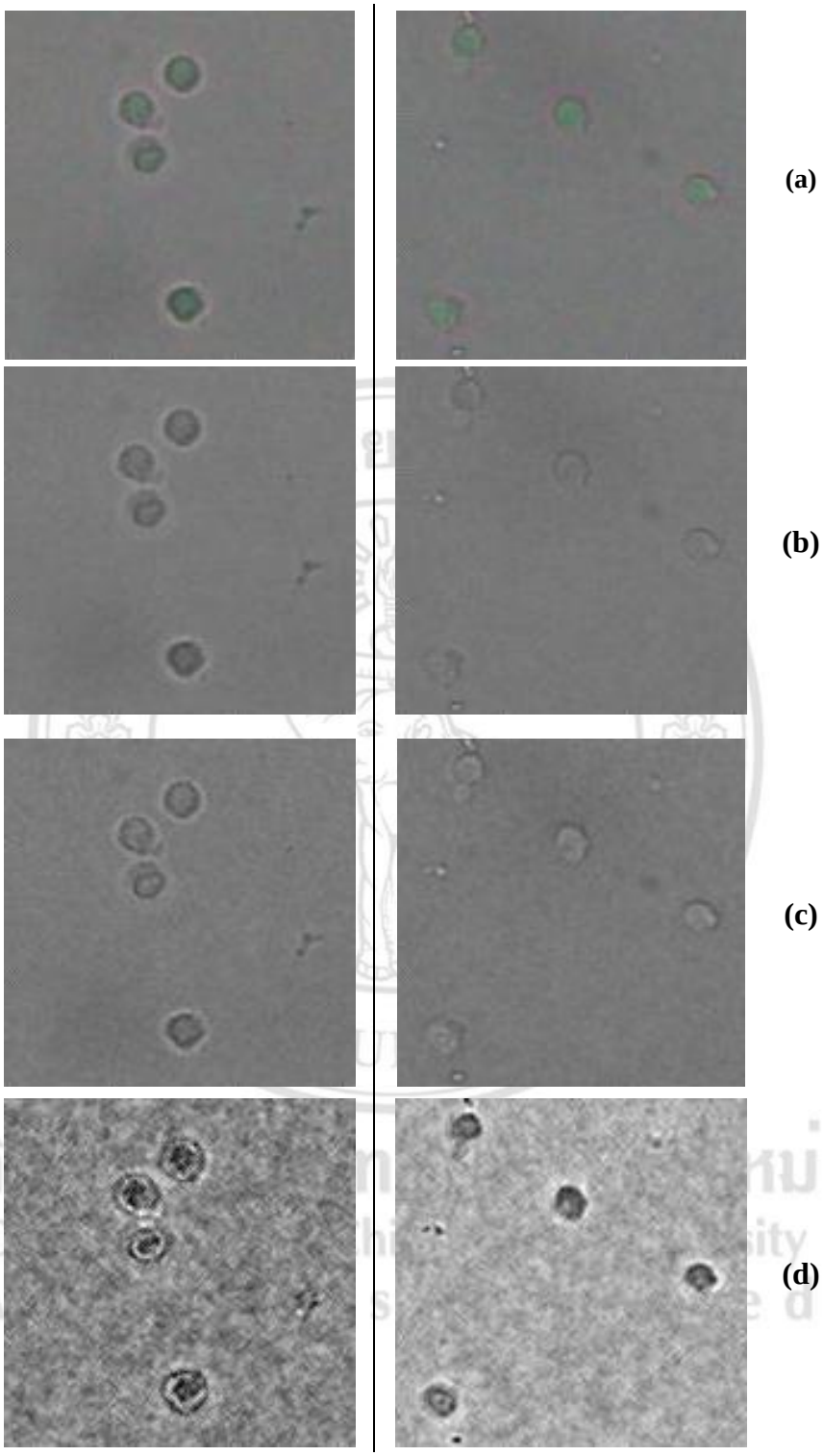


Figure 3.5 Group of (a (left)) clearly-seen and (a (right)) ill-defined boundary cells are represented by (b) gray scale (c) R channel and (d) chrominance images.

2.2) Intensity-Based WBC Candidate Detection: Chrominance Image

The top-hat transform is a good option for dragging cells out of the uneven background based on intensity. This technique is used as the multi-scale morphological filtering to locally enhance the contrast of digital brain MRI images [32]. This is an unsupervised technique that enhance objects which smaller and brighter than a structure element using the morphological opening operation [10]. In contrast, the closing operation intends to enhance the darker objects [32]. We used the closing operation to enhance the interested cells intensities since they are significantly darker than background in chrominance image. To perform the enhancement, gray morphological closing operator is first applied to the image f' (chrominance image) using a small round-shape structure element b_1. The result image h_1 is used as the top of the hat. Then a large round-shape structure element b_2 is used for closing f' to generate the image h_2 as a brim of the hat. Then the enhanced image g is

$$g = h_1 - h_2 \tag{19}$$

where $h_1 = (f' \bullet b_1)$ and $h_2 = (f' \bullet b_2)$. The region of interest is highlighted if it is darker than the background and smaller than b_2. Then the object whose intensity below a threshold t is kept and will be re-segmented using the fuzzy-c mean (FCM) clustering which was explained next.

3) Re-segmentation Using Fuzzy C-Means (FCM) Clustering

Although the WBC candidates detection using multiple gray scale images increase the opportunity to find the WBCs in bright field image, the amount of false detections, such as a background, also increases. As we have known that the refined segmentation is basically used to increase the accuracy of the segmented boundary and the result of algorithm [33]. So, in order to decrease the false detections and increase the correctness of segmentation, a re-segmentation process is generated. We found out that the WBCs can be clearly seen and separated from the background in the RGB image when individually considered in the point of view of a human analyst. In this study, we use the FCM clustering algorithm to distinguish the region of interest

from the background based on the internal cell intensity. In order to perform individual segmentation, we use a binary image g obtained from step 1.2 as the initial region of WBC candidates. Each WBC candidate in RGB image is cropped using a rectangular window 2 pixels larger than its actual diameter along the horizontal and vertical directions. Then the pixel intensities of R, G and B of the cropped image are used as input vectors of the FCM clustering with the parameter m=2, the maximum iteration is 200 with $\varepsilon = 10^{-5}$.

Cluster and Region of Interest (ROI) Selection

After applying FCM clustering, each pixel (x,y) in the FCM clustering's result image F is labeled into n clusters where n = 1 and 2. We assume that the cluster which belongs to the cell content should be located in the middle of the patch image. However, the internal cell intensity is not homogenous. It is possible that the cluster which does not belong to the cell content might appear around the center of F. So, there is a possibility that either cluster number 1 or 2 correspond to WBC, or neither of them. The below algorithm explains the cluster and ROI selection.

1. Get the cluster labeled pixels located around the origin x_0 and y_0 of the image F along both vertical and horizontal directions into the array C whose size is 5x5 pixels. The array C can be denoted by $C = F\left((x_0 - 2 : x_0 + 2),(y_0 - 2 : y_0 + 2)\right)$.

2. Find the number of clusters labeled in C. There are 2 cases possible;

Case 1: Single cluster number in C presents

It is straightforward that the selected cluster should be the present number. Then, we choose the connected component of C as the ROI.

Case 2: More than one clusters in C present

There is a possibility that either cluster 1 or 2 will be chosen. We search for the connected components of each cluster appearing in C. Then which cluster has a larger area is assigned as the ROI.

3. As mentioned, the selected ROI might be either WBC or background. The following steps are applied in order to discard the ROI which probably be a background.

We consider shape and position of ROI in the cropped image. For shape, we computed the proportion of the pixels in the convex hull that are also in the region. The shape of the background obtained from FCM clustering mostly is random, so the given value is low. Finally, we select and assign the ROI as a WBC candidate by considering the following conditions which correspond to the position of ROI;

If the connected component is attached to the image border and the attached length is shorter than half the image length on the corresponding side. Then consider

If the connected component is attached to the image border but on just one side. Then, $H(i,j) = I(x,y)$ where $I(x,y)$ is the chosen region which is projected back to the point (*i, j*) of *H* whose size is similar to the image *g*. The region presented in *g* is assigned as a candidate.

4) Touching and Overlapping Cell Separation

The touching and overlapping cells that appear in our data need to be isolated. The watershed algorithm is popular and capable of dealing with this. This algorithm generates the detached line based on the 3D topography of the image. The 3D topography of the binary image is generated using the Euclidean distance transform in this study. Then the initial marker for the watershed algorithm is indicated by the region minima of the distance map. After that, the watershed algorithm [34] is applied to detach the overlapping cell area.

5) WBC Selection

Due to the ideal shape of WBCs is round, we use the roundness to evaluate the shape of isolated WBC candidates as a criterion for selecting which candidates correspond to WBCs. The roundness of WBC candidates with perimeter *P* and area *A* can be computed using eq.20. Then the candidates are classified as WBC if their roundness is higher than the thresholding value $T_{roundness}$.

$$Roundness = \frac{4\pi A}{P^2} \tag{20}$$

3.3.2. Algorithm for Detecting Positive Cell in Fluorescence Image

As mentioned earlier, green and red fluorescence images are used to identify which WBCs are CD4+ lymphocytes. They can be seen as color spots, meaning "positive", on both fluorescence images. So, the intensity of positive cell is higher than the background. Also, the appearance of positive cell is generally round and rather compact, but the border occasionally blurred due to the diffraction effect of the light. However, some cells have poor appearance and weak intensity. The study [10] described the general three main steps for the fluorescence cell detection approach which consists of noise reduction, signal enhancement and signal thresholding. Our study detected the CD4+ lymphocytes following these general concepts. The detail is described below

1) Noise Reduction and Background Subtraction Using Mathematic Morphological Operation

The median filter is applied to the gray-scaled image *f* for effectively suppressing grain noise on the background. However, the uneven background is still present. So it is necessary to perform image preprocessing. The study [13] estimated the uneven background of the fluorescence images by iteratively fitting the intensity surface with a spline function. However, our study used the gray opening morphological operation for globally background estimation since the implementation is more simply. The background of the grained noise-eliminated image f' can be estimated by applying a dilation operator with a large disk-like structure element *b*. Then the estimated background *h* will be denoted by

$$h = f' - (f' \oplus b) \tag{21}$$

2) Unsupervised Intensity Enhancement

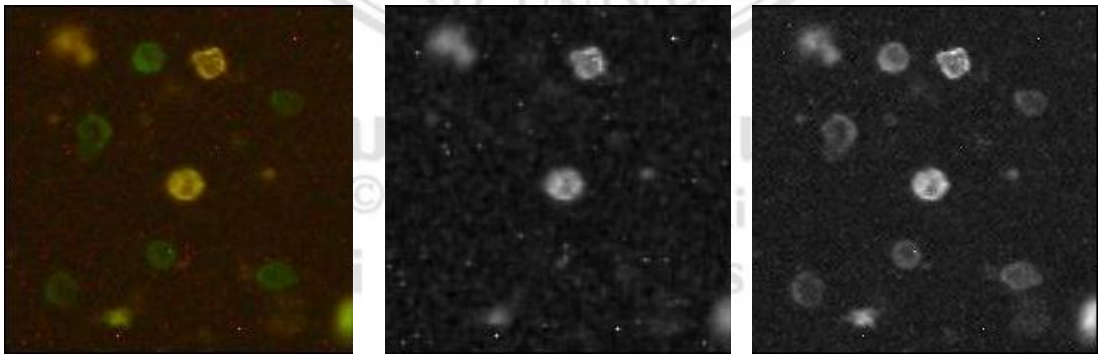
After background subtraction, image intensity enhancement should be performed by employing the multi-scaled top-hat enhancement, which yielded good results for brain MRI images enhancement [32]. Since the intensity of the interested objects is higher than the background, gray opening top-hat operation is needed (unlike the one which was applied to chrominance image in WBCs detection step that

implements the gray closing top-hat). First of all, the opening operator is applied to the background subtracted image *h* for homogenizing the intra-object intensity with the small disk structure element (b_0). The multi-scaled opening top-hat enhancement is shortly described by the following equation

$$H = (H_1) - \frac{1}{n}\sum_{k=1}^{n}(H_2) \quad k = 1,2,...,n \tag{22}$$

where $H_1 = h \circ b_0$ and $H_2 = h \circ b_k$, when b_k is the disk structure element with size b_0 < size b_k < size b_n. Considering the grayscale image for the red fluorescence image, we only use the R channel, which is the best for displaying the red color. On the other hands, cell detection in the green fluorescence image shows a better result if using both R (Fig.3.6b) and G channel (Fig.3.6c) images. Because cells which are positive in both red and green fluorescence images such as CD4+ lymphocytes appear yellow (Fig.11a white arrow). So, the gray opening top-hat operation is applied to both G and R channel images. Then result of green (H_G) and red (H_R) fluorescence images are combined linearly using the following equation

$$H_{combine} = (0.8 \times H_G) + (0.3 \times H_R) \tag{23}$$



(a) **(b)** **(c)**

Figure 3.6 Yellow cells in (a) green fluorescence image are displayed by (b) R channel and (c) G channel.

3.3.3 Find Overlapping Area between Three Segmented Images and Counting

Finally, the result of applying different algorithms to detect the CD4+ lymphocytes in bright field, green and red fluorescence images are 3 binary images denoted by *Br*, *R* and *G*. All obtained binary images are then combined since, as mentioned, CD4+ lymphocytes appear in every image from the stack. The criterion (24) is used to determine the imbrication among three images. It should be noted that the area of detected WBCs in *Br* provides the possible location of CD4+ lymphocytes. Considering *n* connected components in *Br*, suppose $A_{R\cap G}$ is the area of $R\cap G$ and A_{Br} is the area of Br_i. Each connected component *i* in *Br* will be CD4+ lymphocyte if

$$Br_i = \begin{cases} 1 \text{ if } R\cap G\cap Br_i \quad when\ A_{R\cap G} \geq \frac{1}{2}A_{Br} \\ 0 \text{ otherwise} \end{cases} \tag{24}$$

3.4 Performance Evaluation

The results of proposed algorithm were compared to the ground truth drawn by experts. Performance can be evaluated using 2 indices; sensitivity and positive predictive value (PPV). Basically, sensitivity relates to the ability of algorithm to detect cells correctly while the PPV relates to the predicted value of correct detection among the detected cells. Sensitivity and PPV can be denoted by

$$\%Sensitivity = \frac{TP}{TP+FN}\times 100 \tag{25}$$

$$\%PPV = \frac{TP}{TP+FP}\times 100 \tag{26}$$

where *TP* is the correct detection, *FP* is the false detection and *FN* is the missed detection.

The ground truths are generated by using 3 expert's opinion. First expert has the experience on infectious molecular diagnosis and flow cytometer for 3 years, clinical

and microscopic diagnosis for 4 years. Second expert expertize on cell research for more than 5 years. Last expert also has the experience on infectious molecular diagnosis and flow cytometer for 3 years and now studying on stem cell research. It should be noted that all experts are medical technologist.