Identification of Acute Lymphoblastic Leukemia Subtypes in Touching Cells Based on Enhanced Edge Detection

Nenden Siti Fatonah[1,2]*   Handayani Tjandrasa[1]   Chastine Fatichah[1]

[1]*Department of Informatics, Institut Teknologi Sepuluh Nopember, Indonesia*
[2]*Department of Informatics, Mercubuana University, Indonesia*
* Corresponding author's nenden@mercubuana.ac.id

Abstract: Acute lymphoblastic leukemia (ALL) is a type of blood cancer that begins with immature lymphocytes in bone marrow. The challenge in automatic identification of ALL when there are touching cells in image. The previous studies related to the separation of touching cells are still constrained by oversegmentation, undersegmentation, and inaccurate in cell separation. To overcome this problem, this research proposed an algorithm for cell separation based on enhanced edge detection. The objective of this study is identification of L1, L2, and L3, in touching cell based on edge detection and edge linking. Edge detection was performed on grayscale images and Hue images for the area of White Blood Cell segmentation results then the image results were combined. The classification of Acute Lymphoblastic Leukemia (ALL) cells was carried out using the geometry and texture features of cells and nucleus with Support Vector Machine (SVM) as the classifier. The dataset for the training process amounted to 668 ALL single cell images and for the testing process was 301 multi-cell ALL images. The testing results of the identification of ALL subtypes showed that the proposed cell separation method gave accuracy of 75.42%.

Keywords: Acute Lymphoblastic Leukemia, White blood cell, Watershed, Edge detection, Splitting of touching cells, Support vector machine.

1. Introduction

One of the highest causes of death in the world is leukemia [1], caused by abnormal development of white blood cells produced by bone marrow. Classification of white blood cells (WBC) is very necessary to determine the types of leukemia to support the diagnosis of the disease by a doctor. Research on the classification of types of acute leukemia has been widely carried out. However, the presence of WBC touch becomes an obstacle in obtaining accurate single cells in microscopic images of leukemia. An accurate cell separation method will produce better features thereby increasing the accuracy of leukemia cell type classification. Research related to the separation of touching WBC cells has been carried out, in general watershed-based cell separation methods [2, 3], clustering-based methods [4-7], and contour estimation-based [8-15]. In research [2], RGB images were pre-processed by converting to CMY color channel space and using

contrast stretching. Segmentation in research [2] used the Zack algorithm to obtain the threshold and utilized the watershed method for cell separation. Then the classification applied Support Vector Machine (SVM). Then in research [3], cell segmentation was done using the Local density clustering method and connected component analysis, and cell separation was processed by utilizing the Gradient-barrier watershed method. In research [4], the cell preprocessing stage used RGB color space to LAB and the cell segmentation process used the K-Medoids algorithm. Still in research [4] for feature extraction using shape features, visual features, texture features, measurement of pixel values and fractal dimension features. For classification, using 4 classifiers for comparison, namely: Random Forest, Decision Tree, Bayes, and K-Nearest Neighbors (KNN). In research [5], fuzzy morphology method was deployed for segmenting nucleus cells. Separation of touching cells was done by utilizing a

region-based and boundary approach. Research [6] employed RGB image conversion to CIELAB space, carried out the segmentation using the Shadowed C-means (SCM) method, and feature extraction using fractal dimensions, contour signature, shape features, texture features of Gray Level Co-occurrence Matrix (GLCM). Classifiers applied were K-nearest neighbors (KNN), multilayer perceptron (MLP), and support vector machines (SVM). Segmentation using fuzzy C-means method was carried out in research [7], feature extraction used was fractal dimension, contour signature, shape features, color features, texture features of GLCM. SVM was used for classification [7]. Research [8] proposed extracting region of interest (ROI) or cancerous cells employed in two stages for blast cell segmentation. First, all objects in the images were splitted into non-ROI negative class and ROI positive class by utilizing statistical features selected by genetic algorithm. The two classes were identified based on local pixel computation. Secondly, classification of testing data based on the training model of ROI features was carried out by ANN classifier. Research [9] proposed the coarse-to-fine concept, by performing classification among Acute Lymphoblastic Leukemia (ALL), Acute Myeloblastic Leukemia (AML), and healthy cells for the first step. In the second step, sub-types of ALL and AML were identified morphologically based on features of medical knowledge. In the coarse step, a low-pass filter is used for noise removing before extracting nucleus and features which followed by classification in the fine step. In research [10], hybrid model was proposed which combined the segmentation techniques using active contour and fuzzy C means. The segmentation was extracted to obtain the features using statistical and the local directional pattern of chronological sine cosine algorithm. Then, the classification was performed by deep convolutional neural network classifier. Research [11] proposed a computer-aided screening of AML and ALL by analyzing color, morphological and textural features from the blood image. A cell nucleus was segmented using Otsu, and improved using morphological operasions such as opening, hole filling, and boundary cleaning. Geometrical, chromatic, and texture features gave 331 features. The feature optimization is performed by genetic algorithm with support vector machine (SVM) classifier . This research [12] proposed a decision support system for acute leukemia and tested using a public dataset. The system performed K-means clustering for cell segmentation, image refinement, and feature extraction. The cells were classified according to their morphological features. The results were compared with the works of other researchers using the same data. Cell segmentation using fuzzy morphology and fuzzy C-means methods was carried out in research [13], cell extraction was employed using cell area parameters, cytoplasm count, chromatin density, nucleus shape, cytoplasmic form, nucleolus and granules. The classification method using a decision tree was carried out in research [13]. Research [14] converted the RGB color channel space to the gray space and used the threshold with the Otsu's method then the segmentation process by the morphological method, classification using the neural network method. Segmenting cells with the UECS (ultimate erosion convex set) method, then the edge-to-marker and Gaussian association methods on B-splines were proposed to identify and separate cells carried out in the study [15]. Research [16] proposed cell segmentation using the orthogonal gram-schmidt method and the Snakes method. Feature extraction was carried out by employing morphological features, GLCM texture, LBP and Sequential forward selection (SFS) for the feature selection process. And classification using Artificial Neural Networks (ANN), and Support Vector Machines (SVM) methods [16]. In research [17] the first stage is converting the RGB channel space into the HIS channel space, thresholding the S component for the nucleus and Hue for cell identification. Then the process was followed by morphological dilation operation and the median filter. Features were extracted by applying form features, LBP, and classified by using the SVM method [17].

Research [15] proposed the Ultimate Erosion for Convex Sets (UECS) method to identify seed points in a region of interest. White blood cell counts were carried out in research [18] by deploying the Iterative Distance Transform for Convex Sets method to identify seed points in a region of interest. Research [19] employed Fast Radial Symmetry (FRS) to detect objects and used Contour evidence extraction (edge-to-marker association) and Contour estimation (ellipse fitting) to estimate cell contour.

Research [20] applied RGB color space convertion into color space L * a * b, then two color components (a and b) were used as features of the K-means algorithm in segmenting WBC. Research [21] did the segmentation of Acute Lymphocytic Leukemia (ALL) cells using Fuzzy K-means clustering and Automatic Threshold histogram. Research [13] carried out cell segmentation with fuzzy morphology and with fuzzy C-means for cell separation.

Some previous studies related to the separation of touching cells are still constrained by oversegmentation, undersegmentation, and

206

inaccurate in cell separation, so it is still a challenge to develop methods of separating touching cells. The watershed approach allows oversegmentation [2-3], the separation with clustering approach [5-6] is less accurate because it allows the occurrence of areas that are not supposed to appear. Whereas the contour estimation approach using Contour Evidence Extraction with the ellips fitting approach [16] produces contours that do not correspond to the actual cell shape. Research [22] proposed a method of separation in polymer composite particles using an edge detection approach with prewitt operators and obtained good results.

To overcome the problem of separating touching cells such as oversegmentation, undersegmentation, and inaccurate contour, this study aims to use enhanced edge detection approach to separate the touching WBC on the microscopic images of ALL.

This research proposed the method of separating WBC touching cells and estimating single cell contour with edge detection and edge improvement approaches to identify the type of Acute Lymphoblastic Leukemia (ALL). Edge detection is performed on the grayscale image and Hue image based on WBC segmentation results then combines the two edges. Edge detection on the Hue image is to get detailed information on the outer edge of the touching cells, while edge detection on the grayscale image is to get the inside edge details of touching cells. After that, the edge improvement is performed to obtain the accurate contour of each single cell as a result of separating the touching cells.

This study was divided into five stages, namely WBC segmentation, splitting of touching cells, feature extraction, leukemia type classification for each cell, and identification of ALL types in each image. This research extracted the geometry and texture features of the cell and nucleus area in each

single cell then the ALL type classification process was performed using the Support Vector Machine (SVM) classification method. The last stage was counting the number of cells for each type of ALL and the results of identification of ALL types based on the highest number (majority).

This paper is organized as follows. Related work and contribution is shown in introduction in section 1. Section 2 shows methodology of this research. Section 3 shows test results and discussion. Section 4 shows classification results. Section 5 provides conclusion and future work.

2. Methodology

Automated ALL detection in microscopic images requires several stages, namely WBC segmentation, separation of touching cells, feature extraction, classification of ALL type for each cell, and identification of ALL type for each image. The design of the ALL type identification system can be seen in Fig. 1.

The first stage is WBC segmentation which aims to separate the WBC object area and background. The second stage is the separation of touching cells that aims to get a single WBC from several WBC that touch or stick. The results of the second stage are the crops of the single WBC. The next step is feature extraction to get geometry and texture features. Before the feature extraction process, an area segmentation is made which distinguishes the cell and nucleus areas, so that the geometry and texture features are extracted in the cell and nucleus areas.

2.1 Description of the microscopic acute Lymphocytic Leukemia image dataset
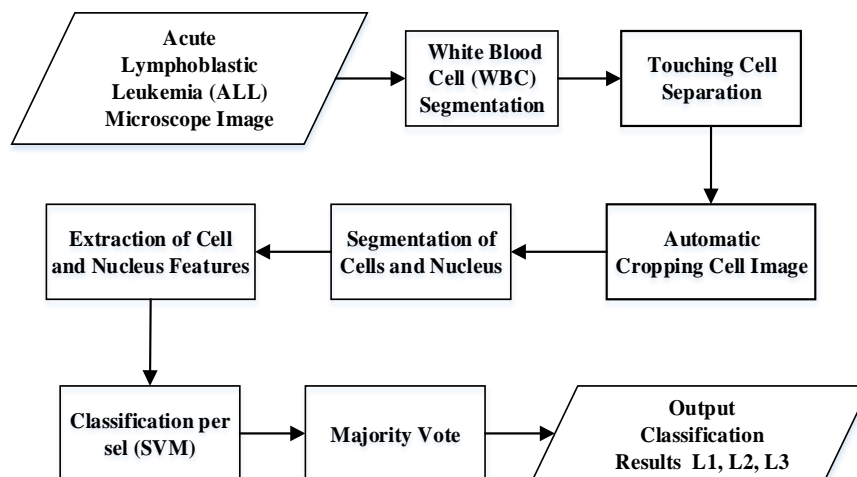
The dataset for the training process amounted to

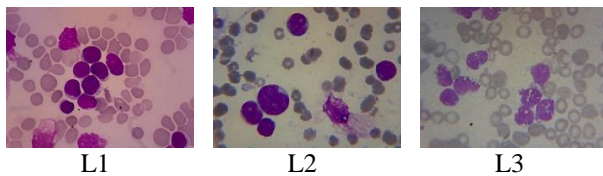Figure. 1 Diagram of type ALL identification system

Figure. 2 Examples of blood smear images of Acute Lymphoblastic Leukemia (ALL)

668 ALL single cell images consisting of 309 L1 typeimages, 174 L2 type images, and 185 L3 type images. The dataset for the testing process was 301 multi-cell ALL images which consisted of 128 L1 type images, 63 L2 type images and 110 L3 type images. The dataset used from dr. Soetomo Hospital in Surabaya. Microscopic images in the dataset used have the same
shooting scale that is 1000x and the same staining process. Microscopic images of the three types of ALL in the dataset used can be seen in Fig. 2.

2.2 White blood cell (WBC) segmentation

WBC segmentation phase aims to get the WBC area and eliminate other components such as red blood cells, platelets, and background. This stage has difficulties if the images in the dataset are used from different staining processes.

The WBC cell segmentation process used an approach like in research [14]. The first process is changing the Leukemia microscopic input image which has an RGB color space to the HSV color space to obtain the Hue component. From the histogram analysis the hue value greater than 0.88 is obtained from the WBC cell area.

The next process is to eliminate the abnormal WBC component (noise), by finding the area of interest (ROI) of objects that have an area of objects (area) of more than 5000 and solidity more than equal to 0.50. The final process is to remove the truncated cell object that is on the edge of the image if present.

2.3 Splitting touching WBC

This stage is to do the separation of touching WBC and get the results of single cell cropping from the results of separation. From the results of the single cell cropping, feature extraction will be performed. There are three methods for splitting touching WBC that used in this research i.e. watershed, clustering, and edge detection.

2.3.1. Watershed method

Watershed transformation is to treat a grayscale image as a topographic map, with the intensity representing its height. The concept of watershed is to fill up all catchment basins with water starting with the lowest catchment basin until all the catchment basins begin to be flooded with water. Watersheds are built as different basins initiate to unite that create lines on ridge tops. The process is ended when the water reaches the maximum peak of every relief [23].

2.3.2. Clustering method

The clustering method used for separating touching cells in this study is based on searching for cell markers first using the IDTCS method [14] which is used as the centroid of each cell, then grouping pixels in each cell to the closest centroid. The IDTCS approach iterates the distance transform procedure for all objects. The cell is flagged when the concavity size is less than a threshold value.

2.3.3. Proposed method

The proposed method for splitting touching WBC using enhanced edge detection combined edge detection and edge linking. The flowchart of proposed method for the separation of cells in touching WBC can be seen in Fig. 3 and the pseudocode in Fig. 4.

The inputs at the stage of touching cell separation is a binary image from the WBC segmentation image to get the WBC area, and the original microscopic ALL image to get grayscale and HSV information from each WBC area. The edge detection process is then performed on the grayscale image and on the Hue component image in the WBC segmentation area result. The edge detection method used is the Canny method. The next process is to merge the two images resulting from edge detection of grayscale images and Hue images. Edge detection in the Hue image is to obtain detailed information on the outer edge of the touching cells while edge detection in the Grayscale image is to obtain the inner edge details of touching cells. After that an edge improvement is performed to obtain the contour of each single cell as a result of separating the touching cells. Edge improvement consists of two processes, namely the process of connecting pixels that are still broken using the gapfilling function and the second process is filtering using the Gaussian method. The next step is to imfill the edges of the image. Then to get the separated single cell from the touching WBC, the subtraction process is done between the image of the imfill
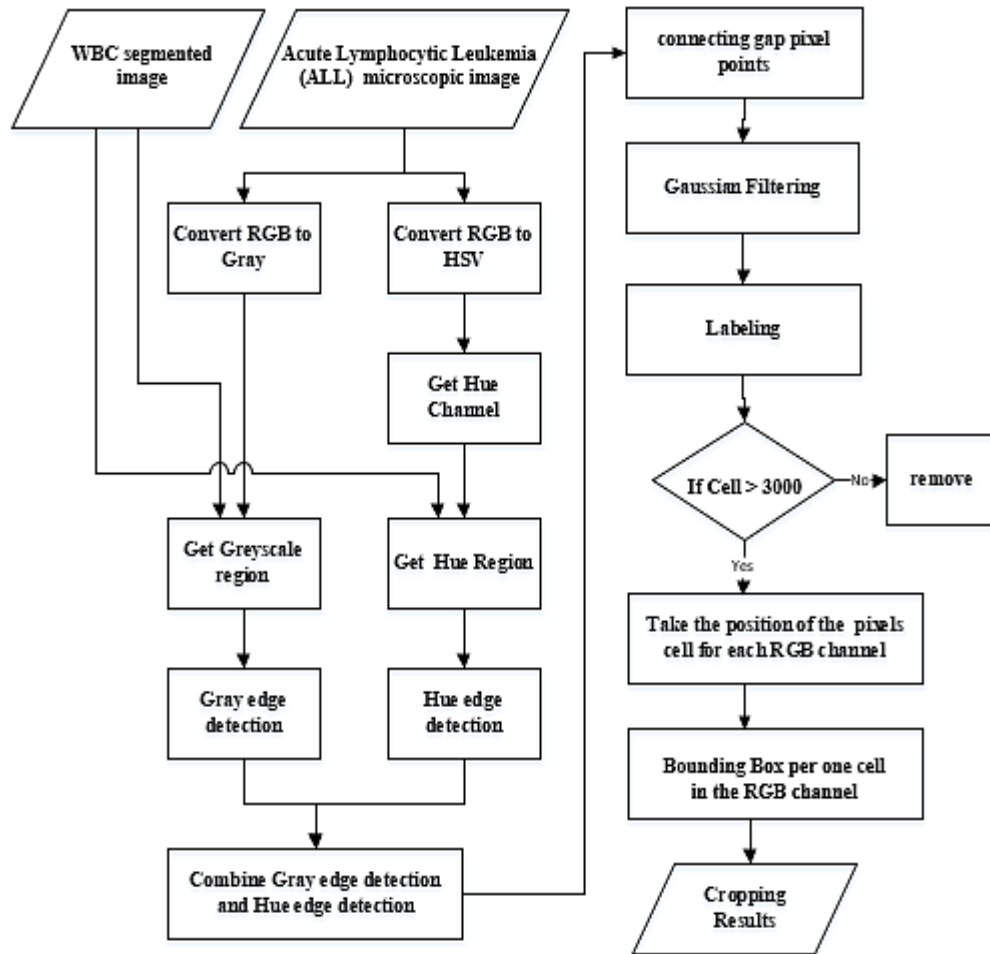
Figure. 4 Cell splitting and cropping flowchart

process and the image of the edge result. If the area of a single cell is less than 3000 then the cell area is removed. The final stage is the process of cropping the WBC single area of the original RGB image using the boundingbox approach so that it produces as many RGB images as the single WBC on the original image. The output of this process is cropped RGB cell images. The pseudocode of the proposed method for separation of touching cells can be seen in Fig. 4.

2.4 Feature extraction on each WBC

2.4.1. Nucleus and cell segmentation of each WBC

Before performing the feature extraction process for each WBC cell, it is necessary to do cell segmentation to obtain the nucleus and whole cell area of each WBC from the ALL microscopic image. The stages of this cell segmentation can be seen in Fig. 5.

The input is the cropped RGB image for every WBC, then the thresholding process is done using the

Input: Binary segmented image B and Original image I

Output: Contour extraction of touching WBC image S

Algorithm:
1. G = Convert I to grayscale image
2. H = Convert I to Hue image
3. $G1$ = Take grayscale value from G for the area of segmented WBC B
4. $H1$ = Take Hue value from H for the area of segmented WBC B
5. $G2$ = Canny edge detection ($G1$)
6. $H2$ = Canny edge detection ($H1$)
7. $U = G2 + H2$
8. C = Connecting gap pixels of U by adding 19 pixels.
9. F = Gaussian Low Pass Filter (C)
10. K = $binarization\ F$
11. L = imfill (K)
12. $S = L - K$

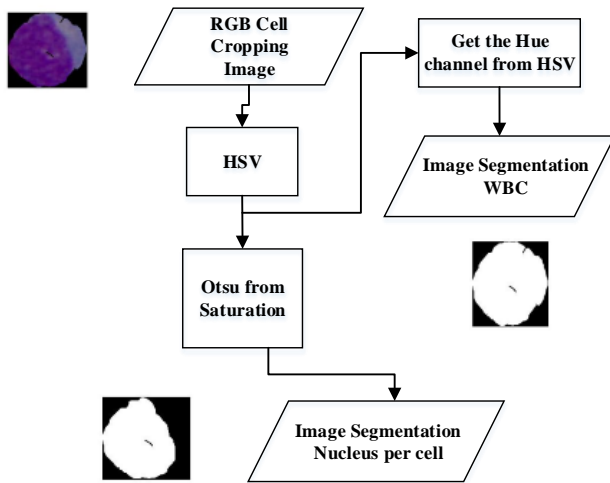Figure. 3 Pseudocode of touching cell separation based on the enhanced edge detection

Figure. 5 Nucleus and cell segmentation for image per cell

Otsu method of Saturation color space on the cropped RGB image. The result of segmentation is a binary image with the nucleus area. To get the cell area, the binarization is carried out in the cropped RGB image to get the whole cell area then the binary image of the cell is subtracted by the binary nucleus image to get the cytoplasmic area.

2.4.2. Geometry and texture feature extraction in nucleus and cell area

The features used in this study are composed of 34 features as in the study [4] with each divided into features for white blood cells and nucleus. The features used are the geometry and texture features as follows:

Area: Number of pixels in a region or object.
Perimeter: Number of pixels on the boundary or edge of a region or object.
Roundness: A measure of how close the shape of an object is to the shape of a circle as in Eq. (1).

$$Roundness(R) = 4.pi.A(R)/P^2(R) \qquad (1)$$

where A is the area, P is the perimeter.

Eccentricity: A measure of how big the difference in the shape of an object is from the shape of a circle, as in Eq. (2).

$$Eccentricity = \frac{Distance\ between\ ellipse\ Foci}{Length\ of\ Major\ axis} \qquad (2)$$

Major axis is the longest axis of an ellipse.

Solidity: The ratio between the area and the convex area as in Eq. (3).

$$Solidity = \frac{A(R)}{Convex\ area} \qquad (3)$$

Equivalent Diameter: The value is defined as in Eq. (4).

$$Equivalent\ diameter = 2.\sqrt{\frac{A(R)}{pi}} \qquad (4)$$

Orientation: the angle between the horizontal axis and the ellipse main axis. The main axis line goes through the centroid and along the region longest part.
Radius: the radius of a region or object calculated according to Eq. (5).

$$Radius = Equivalent\ diameter/2 \qquad (5)$$

Centroid is a region center, an arithmetic average of all (x, y) coordinates in the region.

N / C ratio: determines the ratio of the area of the nucleus and whole cells as in Eq. (6).

$$\frac{N}{C}ratio = \frac{Area\ of\ the\ nucleus}{Area\ of\ cell} \qquad (6)$$

The texture features can be obtained from the features of Gray Level Co-occurrence Matrix (GLCM) or Local Binary Pattern (LBP) methods in processing image data [24][25]. The texture features in this study use the Gray Level Co-occurrence Matrix (GLCM) descriptor as in Eqs. (7) - (10).

$$Contrast = \sum_{i,j}|i-j|^2 P_{ij} \qquad (7)$$

P_{ij} is the element (i,j) of the normalized GLCM matrix.
Correlation: the correlation of a pixel with the neighbors over the whole image.

$$Correlation = \sum_{i,j-0}^{N-1} P_{ij}\frac{(i-\mu)(j-\mu)}{\sigma^2} \qquad (8)$$

where:
P_{ij} is the element (i,j) of the normalized symmetrical GLCM.
N is the number of gray levels in the image, μ is the GLCM mean, as in Eq. (9):

$$\mu = \sum_{i,j-0}^{N-1} iP_{ij} \qquad (9)$$

σ^2 is the GLCM variance, as in Eq. (10):

$$\sigma 2 = \sum_{i,j=0}^{N-1} P_{ij}(i-\mu)^2 \qquad (10)$$

Energy: Determines the uniformity of the image.

$$Energy = \sum_{i,j}(P_{ij})^2 \qquad (11)$$

Homogenity: Measuring the proximity of the GLCM element distribution to the diagonal.

$$Homogeneity = \sum_{i,j} \frac{P_{ij}}{1+|i-j|} \qquad (12)$$

Euler is counting the number of holes in leukemia cells to detect vacuola.
Maximum intensity is the highest pixel value in the region.
Minimum Intensity is the lowest pixel value in the region.
Average intensity is the average pixel value in the region.

2.5 Classification of ALL subtypes for each WBC

The stages for classifying ALL subtypes in each WBC can be seen in Fig. 6. In this study, Support Vector Machine (SVM) is used in the classification of Acute Lymphocytic Leukemia (ALL) subtypes. SVM is a prediction method for classification and regression. The concept of SVM is to find the optimal separator function that can separate data from two classes, latter on SVM is developed for multiclass classification.

Training data is used as the input for the SVM training process to produce the SVM model. The testing process is carried out on ALL microscopic images after going through several stages such as WBC segmentation, cell separation, and feature
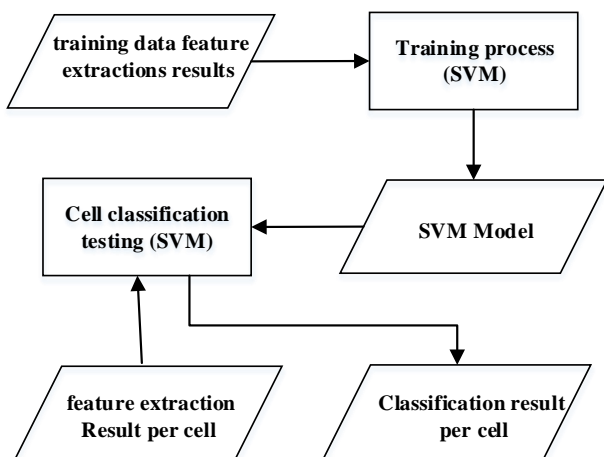

Figure. 6 Stages of classification

extraction of each WBC cell. From the result of the feature extraction, each WBC cell is predicted based on the SVM model to produce an ALL subtype, namely L1, L2, or L3. The number of prediction results of ALL subtypes is the same as the number of WBC in an ALL microscopic input image.

2.6 Identify the ALL subtype of each image

Determination of the type of ALL in each leukemia microscopic image is done with the concept of Majority Vote according to the number of WBC cell classification results in the image. If in the image the classification results are more cells of L1 class then the input image is of L1 type, although there are some cells whose prediction results are not in L1 class.

3. Test results and discussion

3.1 White blood cell segmentation test results on ALL images

The first stage in this study is the segmentation of white blood cells (WBC) on the microscopic image of Acute Lymphocytic Leukemia (ALL). The output of this segmentation stage is the image of white blood cell segmentation.

The output stage of the WBC segmentation process on ALL microscopic images can be seen in Fig. 7. Fig. 7 (a) shows the original image and Fig. 7 (b) gives the image of binarization. Fig. 7 (c) shows the Hue color channel image obtained from the binary image and the image in Fig. 7 (d) is from the hue channel image after thresholding performed to get the White Blood Cell and remove the Red Blood Cell image, but the results are not clean and the noise still appears and then in Fig. 7 (e) the WBC image is obtained after cleaning the WBC area from the small objects and noise that is still visible in Fig. 7 (d).

3.2 Test results with the proposed separation method

3.2.1. The cell separation results of the proposed separation method

The stages of WBC separation on ALL microscopic images is exemplified in Fig. 8. Fig. 8 (a) shows WBC segmentation, then Fig. 8 (b) and 8 (c) give the results of the canny process in the hue image and grayscale image respectively. The image in Fig. 8 (d) is the summation of images in Fig. 8 (b) and 8 (c). The following processes are filling the gap of less than 19 pixels and linking the edges, Gaussian
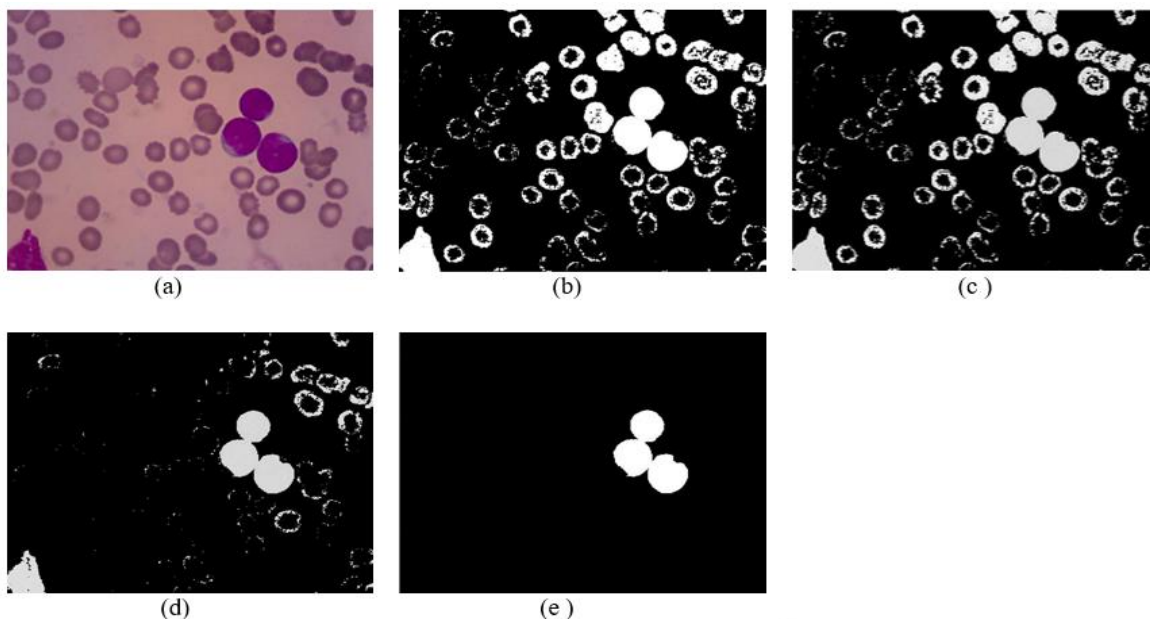
Figure. 7 An example of WBC segmentation in an ALL image: (a) RGB Image, (b) Binary image,
(c) Hue channel image, (d) White Blood Cells noise, and (e) White Blood Cells segmentation.
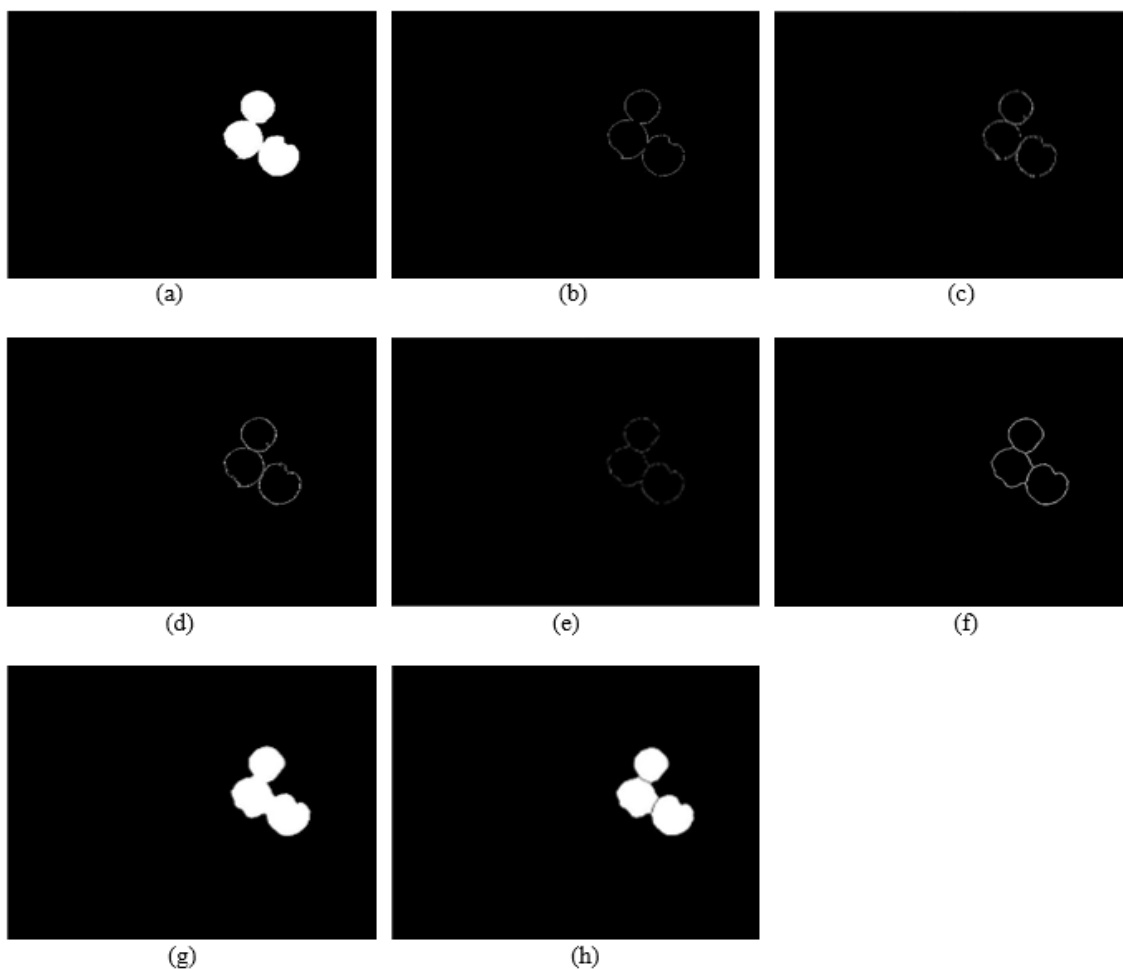

Figure. 8 The stages of White Blood Cell separation in an ALL image: (a) White Blood Cell segmentation, (b) Canny of
Hue channel, (c) Canny of Gray channel, (d) Combined result of b and c, (e) Image after Gap filling, (f) Filtered image,
(g) Hole filled image, and (h) Image of splitted cell

low pass filtering, and binarization as the results shown in Fig. 8 (e) and 8 (f). Furthermore, the ouput of the hole filling process is shown in Fig. 8 (g). The separated cells shown in Fig. 8 (h) are obtained by subtracting the image in Fig. 8 (g) by the image in Fig. 8 (f).

Following is the comparison of the proposed method for cell separation using the enhanced edge detection, the IDTCS clustering method, and the Watershed method, as shown in Fig. 9. IDTCS clustering uses the IDTCS method [14] to get

centroids then grouping pixels according with the obtained centroids. The results of segmentation in Fig. 9, based on the three methods: proposed method, watershed and clustering. The segmentation results from clustering (a2-c2) and watershed (a3-c3) are oversegmented, while the results of the proposed method (a1-c1) are better than the results of watershed or clustering. Fig 9. (a1 - c1) show that the proposed method can precisely separate each touching cells. In contrast to other methods,
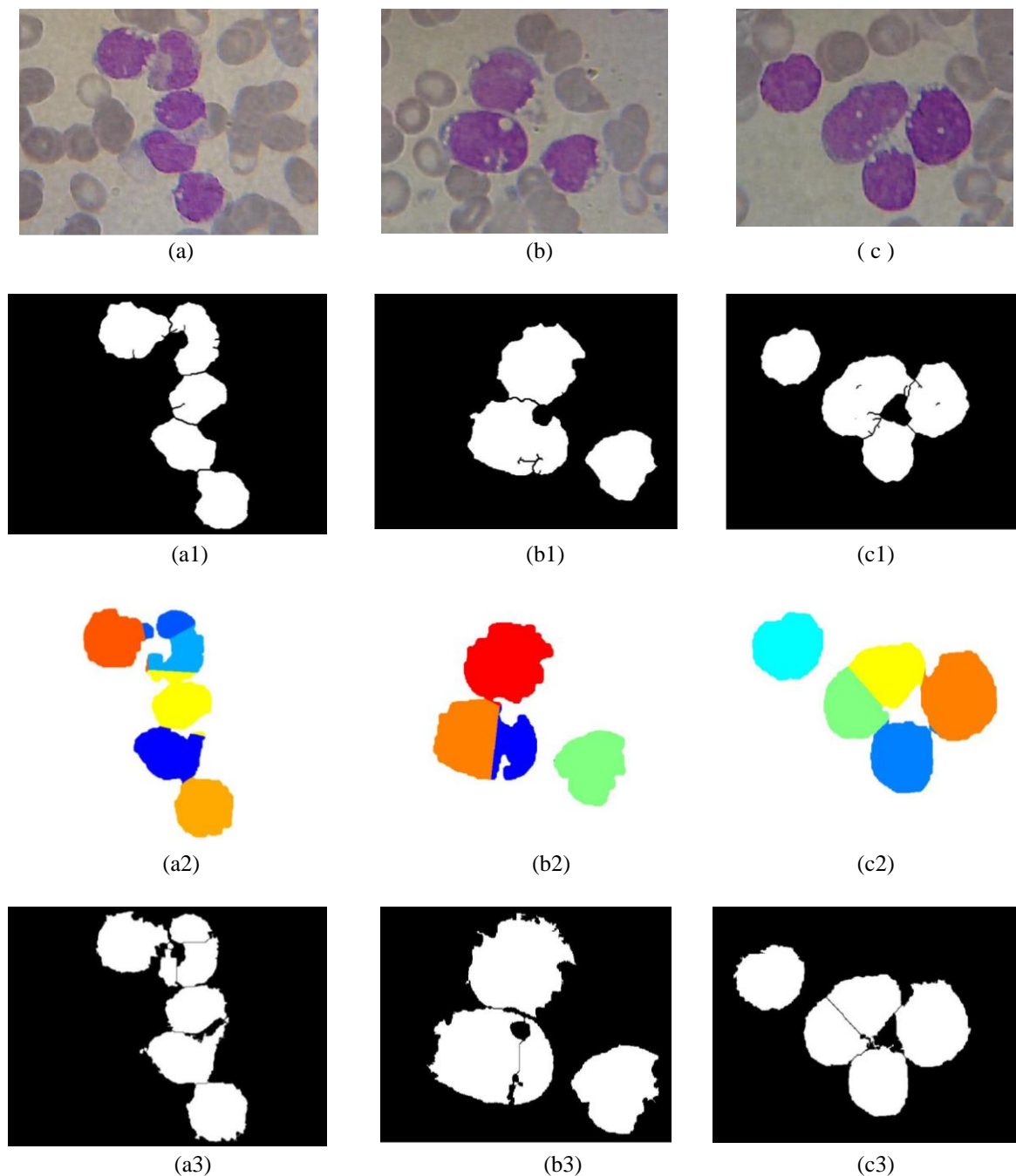

Figure. 9 (a-c) Original Image, (a1-c1) results of enhanced edge detection method (proposed method), (a2-c2) results of IDTCS clustering method, and (a3-c3) results of Watershed method

213

Figure. 10 Results of single WBC cropping using the
proposed separation method

Table 1. Classification sensitivity results of 3 methods of
white blood cell separation in ALL images

ALL Type	Propose Method (%)	Watershed (%)	IDTCS clustering (%)
L1	90.63	79.69	80.47
L2	77.78	85.71	82.54
L3	56.36	43.64	14.55

Table 2. Confusion matrix result of ALL subtype
identification using the proposed method

		Predicted			
		L1	L2	L3	FN
Actual	L1	116	12	0	12
	L2	13	49	1	14
	L3	45	3	62	48
	FP	58	15	1	

oversegmentation occurs for the clustering method as shown in Fig. 9 (a2 - c2) due to misdetection of centroids and inaccuracy of determining boundaries between cells. Oversegmentation also occurs for watershed method as in Fig. 9 (a3 - c3) when there is a vacuola in the middle of the cell.

3.2.2. WBC single cropping results with the proposed separation method

After separating the white blood cells based on the enhanced edge detection, the result is used as the mask for the RBG image, each cell in the RGB image is cropped individually using the bounding box. The examples of the cropped cell are shown in Fig. 10. Then, each cropped cell is processed to extract the features that will be tested by the previously trained SVM model (see Fig. 6). The identification of L1, L2, and L3 classes for an ALL image is determined by the majority vote of WBC cell classification results in the image.

4. Classification results

4.1 Classification sensitivity for different methods of touching cell separation

The sensitivity results for the classification of ALL subtypes by the method of separating the touching cells using Watershed, IDTCS clustering, and the proposed method can be seen in Table 1.

It can be seen in Table 1 that for ALL subtype of L1 the percentage of sensitity for the Watershed method is 79.69%, for the proposed method is 90.63%, and IDTCS clustering is 80.47%, meaning that the proposed method is better than the Watershed method, and IDTCS clustering. For L2 subtype the sensitity results for the Watershed method are 85.71%, for the proposed method is 77.78%, and IDTCS clustering is 82.54%, meaning the watershed method is better than the proposed method and the IDTCS clustering method. For L3 subtype the sensitity of the Watershed method is 43.64%, and for the proposed method is 56.36%, and IDTCS clustering is 14.55%, meaning that the proposed method is better than the Watershed method and the IDTCS clustering method.

4.2 Confusion matrix of the proposed method

The result of confusion matrix for identification of ALL subtypes based on the proposed cell separation method can be seen in Table 2.

The examples of the cropped cell are shown in Fig. 10. Then, each cropped cell is processed to extract the features that will be tested by the previously trained SVM model (see Fig. 6). The identification of L1, L2, and L3 classes for an ALL image is determined by the majority vote of WBC cell classification results in the image.

The classification results for the proposed method for ALL L1 types identified as L1 of 128 images are 116 images and those incorrectly identified as L2 are 12 images. Here it is seen that the resulting method is accurate for identifying type of L1. In the ALL L2 type, there are 49 images correctly identified as L2 out of 63 images and 13 images incorrectly identified as L1 and 1 image incorrectly identified as L3. In conclusion, there are many types of ALL L2 that are detected as L1 because L2 and L1 have similar characteristics. In the ALL L3 type, from 110 images there are 62 images correctly identified, 45 images incorrectly identified as L1, and 3 images incorrectly identified as L2. The ALL L3 type that identified as L1 is 45 images because there are many L3 type images that have several WBC that are very similar to L1. The accuracy of all images is 75.42%.

5. Conclusion

In this paper, we propose methods for separating touching cell and contour estimation based on edge detection and edge improvement. With better touching cell separation, the extraction of individual WBC gives good features, so that the ALL subtype can be identified more accurate. From the test results

214

of Acute Lymphoblastic Leukemia (ALL) microscopic image dataset, for L1 class the proposed separation method has higher sensitivity compared to the Watershed method and the IDTCS clustering method. For the L1 type the proposed method gets sensitivity of 90.63%, the IDTCS clustering method and the Watershed method has sensitivity of 80.47% and 79.69% respectively. For L2 type, Watershed method has higher sensitivity compared to the proposed method and IDTCS clustering. Watershed method gives sensitivity of 85.71%, IDTCS clustering method and the proposed method gives sensitivity of 82.54% and 77.78% respectively.

For L3 types, the proposed method has higher sensitivity compared to the Watershed and IDTCS clustering methods. For L3 class, the proposed method obtains sensitivity of 56.36%, the Watershed method and the IDTCS clustering method has sensitivity of 43.64% and 14.55% respectively. The accurate identification of ALL is required to develop in the clinical application, so it is still a challenge to find the best method of each process especially segmentation and feature extraction. In addition, the future research can be extended to classify other leukemia types such as Acute Myeloblastic Leukemia (AML).

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization, Nenden Siti Fatonah, Handayani Tjandrasa, Chastine Fatichah; methodology, Nenden Siti Fatonah, Handayani Tjandrasa, Chastine Fatichah; software, Nenden Siti Fatonah; validation, Nenden Siti Fatonah, Handayani Tjandrasa, Chastine Fatichah; formal analysis, Nenden Siti Fatonah, Handayani Tjandrasa, Chastine Fatichah; investigation, Nenden Siti Fatonah, Handayani Tjandrasa, Chastine Fatichah; resources, Nenden Siti Fatonah, Handayani Tjandrasa, Chastine Fatichah; data curation, Nenden Siti Fatonah, Handayani Tjandrasa, Chastine Fatichah; writing—original draft preparation, Nenden Siti Fatonah; writing—review and editing; Handayani Tjandrasa, Chastine Fatichah; visualization, Handayani Tjandrasa, Chastine Fatichah; supervision, Handayani Tjandrasa, Chastine Fatichah.

Acknowledgments

References

[1] National Cancer Institute, "What You Need To Know About Leukemia", *U.S. Department of Health and Human Services*, 2013. [Online]. Available: http://docplayer.net/32703074-What-you-need-to-know-about-leukemia.html.

[2] L. Putzu, G. Caocci, and C. Di, "Artificial Intelligence in Medicine Leucocyte classification for leukaemia detection using image processing techniques", *Artif. Intell. Med.*, Vol. 62, No. 3, pp. 179–191, 2014.

[3] H. Yang and N. Ahuja, "Automatic segmentation of granular objects in images: Combining local density clustering and gradient-barrier watershed", *Pattern Recognit.*, Vol. 47, No. 6, pp. 2266–2279, 2014.

[4] V. Acharya and P. Kumar, "Detection of acute lymphoblastic leukemia using image segmentation and data mining algorithms", *Med. Biol. Eng. Comput.*, Vol. 57, No. 8, pp. 1783–1811, 2019.

[5] C. Fatichah, D. Purwitasari, V. Hariadi, and F. Effendy, "Overlapping White Blood Cell Segmentation and Overlapping White Blood Cell Segmentation and Counting on Microscopic Blood Cell Images", Vol. 7, No. 3, pp. 1271–1286, 2014.

[6] S. Mohapatra, D. Patra, and S. Satpathy, "An ensemble classifier system for early diagnosis of acute lymphoblastic leukemia in blood microscopic images", *Neural Comput. Appl.*, Vol. 24, No. 7–8, pp. 1887–1904, 2014.

[7] S. Mohapatra, D. Patra, and S. Satpathi, "Image analysis of blood microscopic images for acute leukemia detection", In: *Proc. of 2010 Int. Conf. Ind. Electron. Control Robot. IECR 2010*, pp. 215–219, 2010.

[8] S. S. Al-jaboriy, N. N. A. Sjarif, S. Chuprat, and W. M. Abduallah, "Acute lymphoblastic leukemia segmentation using local pixel information", *Pattern Recognit. Lett.*, Vol. 125, pp. 85–90, 2019.

[9] J. Laosai and K. Chamnongthai, "Classification of acute leukemia using medical-knowledge-based morphology and CD marker", *Biomed. Signal Process. Control*, Vol. 44, pp. 127–137, 2018.

[10] K. K. Jha and H. S. Dutta, "Mutual Information based hybrid model and deep learning for Acute Lymphocytic Leukemia detection in single cell blood smear images", *Comput. Methods*

Programs Biomed., Vol. 179, p. 104987, 2019.

[11] J. Rawat, A. Singh, B. HS, J. Virmani, and J. S. Devgun, "Computer assisted classification framework for prediction of acute lymphoblastic and acute myeloblastic leukemia", *Biocybern. Biomed. Eng.*, Vol. 37, No. 4, pp. 637–654, 2017.

[12] A. S. Negm, O. A. Hassan, and A. H. Kandil, "ORIGINAL ARTICLE A decision support system for Acute Leukaemia classification based on digital microscopic images", *Alexandria Eng. J.*, Vol. 57, No. 4, pp. 2319–2332, 2018.

[13] C. Fatichah, M. L. Tangel, F. Yan, J. P. Betancourt, M. R. Widyanto, and F. Dong, "Fuzzy feature representation for white blood cell differential counting in acute leukemia diagnosis", *Int. J. Control. Autom. Syst.*, Vol. 13, No. 3, pp. 742–752, 2015.

[14] S. Nazlibilek, D. Karacor, T. Ercan, M. H. Sazli, O. Kalender, and Y. Ege, "Automatic segmentation, counting, size determination and classification of white blood cells", *Meas. J. Int. Meas. Confed.*, Vol. 55, pp. 58–65, 2014.

[15] C. Park, J. Z. Huang, J. Ji, and Y. Ding, "Segmentation, Inference and Classification of Partially Overlapping Nanoparticles", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 35, No. 3, pp. 1–1, 2012.

[16] S. H. Rezatofighi and H. Soltanian-Zadeh, "Automatic recognition of five types of white blood cells in peripheral blood", *Comput. Med. Imaging Graph*, Vol. 35, No. 4, pp. 333–343, 2011.

[17] V. Singhal and P. Singh, "Local Binary Pattern for automatic detection of Acute Lymphoblastic Leukemia", *Commun. (NCC), 2014 Twent. Natl. Conf.*, pp. 1–5, 2014.

[18] N. S. Fatonah, H. Tjandrasa, and C. Fatichah, "Automatic Leukemia Cell Counting using Iterative Distance Transform for Convex Sets", *Int. J. Electr. Comput. Eng.*, Vol. 8, No. 3, p. 1731, 2018.

[19] S. Zafari, "Segmentation of Overlapping Convex Objects", *Master's Thesis*, August, 2014.

[20] N. M. Salem, "Segmentation of white blood cells from microscopic images using K-means clustering", In: *Proc. of Radio Science Conference (NRSC), 2014 31st National*, pp. 371–376, 2014.

[21] F. Scotti, "Robust segmentation and measurements techniques of white cells in blood microscope images", *Conf. Rec. - IEEE Instrum. Meas. Technol. Conf.*, pp. 43–48, 2006.

[22] F. A. Hermawati, I. M. Kastiawan, and Muhyin, "Segmentation method for splitting filler particle of polymer composite in microstructure image", In: *Proc. of the 2nd International Conference on Vacational Innovation and Applied Science (ICVIAS 2019)*, 2019 .

[23] A. S. Kornilov and I. V. Safonov, "An overview of watershed algorithm implementations in open source libraries", *J. Imaging*, Vol. 4, No. 10, 2018.

[24] I. Nurhaida, H. Wei, R. A. M. Zen, R. Manurung, and A. M. Arymurthy, "Texture fusion for batik motif retrieval system", *Int. J. Electr. Comput. Eng.*, Vol. 6, No. 6, pp. 3174–3187, 2016.

[25] H. Harwikarya, "Comparison of Texture's Features of SAR Images Based on Grey Level Co-occurrence Matrix : Case on Multi Polarization P Band and L Band", *Int. Res. J. Comput. Sci.*, No.3, March, 2016.