Color-Based Hybrid Modeling to Classify the Acute Lymphoblastic Leukemia

Arif Muntasa[1]   Muhammad Yusuf[1]*

[1]*Informatics Engineering Department Universitas Trunojoyo Madura, Indonesia*
* Corresponding author's Email: muhammadyusuf@trunojoyo.ac.id

Abstract: In this research, a new approach color based hybrid modeling to separate and classify the Acute Lymphoblastic Leukemia is proposed to resolve the classification tasks. A series of the process is offered to conduct the classification tasks, i.e., Pre-processing, Image Segmentation, Features Extraction, and Similarity Measurement. Furthermore, the Otsu thresholding-based contrast stretching model is proposed to enhance the image quality on the Pre-processing stage. Moreover, the largest object selection of the found objects indicates the right choice to capture the desired object. Furthermore, the feature extraction is performed by combining the color based densitometry and shape features. Lastly, the feature extraction results are measured by the Euclidian Distance and Manhattan method to classify the sample used, which is Acute Lymphoblastic Leukemia Image Database (ALL-IDB). The results of the proposed model have produced 95.38% accuracy. Therefore, it showed that the classification had produced higher accuracy than the others, i.e., Naïve Bayesian, Color Correlation, Fuzzy-based Leukemia Detection, Hausdrof SVM-based Leukemia Detection, and Automated Differential (Learning Vector Quantization, Multi-Layer Perceptron, and Support Vector Machine.

Keywords: Segmentation**,** Classification**,** Color-based hybrid**,** Acute Lymphoblastic Leukemia.

1. Introduction

The development of computer science research has been widely expanded to solve many problems in biomedical science, engineering, and imaging research fields. A collaboration between the theoretical and applied computer science has inspired many researchers to improve research methods and results. The research results have a contribution in the biomedical fields, i.e., blood vessel segmentation [1–3], optic disc detection [4–7], diabetic retinopathy detection [8, 9], and even cancer detection [10, 11].

Acute lymphoblastic leukemia (leukemia) is a type of white blood cell cancer. It is categorized as the most dangerous disease type of cancer that attacks a person's immune system. Many researchers have tried to detect the disease early, unfortunately, it requires high cost and takes a long time to classify the disease.

Complete blood cell count, as well known as CBC, is carried out to detect leukemia before the biomedical imaging research field, is developed by many researchers. This approach is conducted by the Immunophenotyping model, where flow cytometer counts the blood sample. Furthermore, this sample is discarded after processing. This model is not adequate to detect leukemia. One of the efforts is to research in the imaging field. Along with the development of research results in the field of image processing, some approaches have been proposed and produced by many researchers to identify and classify leukemia [12–24].

One of the most popular methods that have been published to classify is the Naïve Bayes. This classifier is performed to classify and detect leukemia. The Naïve Bayes has approached for classification based on the maximum probability of the object characteristic. The distribution of the image pixels influences the characteristics of the object, including the average and variance of the purpose [25]. Unfortunately, the process carried out is not preceded by image enhancement. The segmentation generates inaccurate results, so the extraction results also do not reflect the overall characteristics of the object. Therefore, the classification has produced lower

accuracy. Some similar researches have also been conducted by several researchers, where the segmentation and feature extraction were performed without pre-processing [26–28]. However, pre-processing contributes to increasing accuracy results.

Other approaches to classify leukemia has been produced by several researchers, i.e., Watershed technique [29], Fuzzy-based Leukemia Detection Clustering [30], and Color Correlation [13]. The research results of Fuzzy-based Leukemia Detection produces 93% accuracy, even though the classifier is a support vector machine method [30]. This method carried out the transformation from image color into L*a*b* as image enhancement. This transformation will cause some loss of image information so that the segmentation results become invalid. The results will produce feature extraction that is not representing the real object. This weakness can be improved by our proposed method, where pre-processing is carried out by a series of processing, i.e., channel separation, green channel adjustment, re-arrange image channel, and image mapping to the HSV model. This is a new aim approach to obtain the best image results before the segmentation process.

A disadvantage of the Fuzzy-based Leukemia Detection is the object class determination based on the distance of the object, while the image pixel distribution is not considered for mapping the object class. This rule will generate an inaccurate object. To overcome the weakness, we proposed the binary segmentation by viewing of the image pixels, where the threshold value can be adaptively determined by the Otsu method. There are three stages of the segmentation, i.e., binary image transformation, the biggest object area selection, and re-arrangement of the source and the negative of the binary image result transformation.

The segmentation results will be further processed to gain the feature of the object. To acquire the principal features of the object, we proposed a hybrid model, i.e., the shape and color based densitometry. The shape represents an object's appearance, i.e., the object area, perimeter, and circularity features. Color-based densitometry describes energy, entropy, correlation, and variance. The images used to evaluate the algorithm are a color model, but several researchers have transformed the original image into grayscale before feature extraction [31–33].

Furthermore, [34] employed the Salp Swarm integrated Dolphin Echolocation-based Support Vector Neural Network (SSDE-SVNN) classifier. This research utilized ALL-IDB1 contains 108 images with the existence of 39000 blood details. Additionally, the expert oncologists have labeled the lymphocytes. The results achieve 0.97 of accuracy, 0.97 of specificity, and 1 of sensitivity, respectively. Also, the Mean Square Error (MSE) is 0.1272. The proposed method has better results than GFNB, ELM, KNN, SVM, and Naive Bayes. Besides, the processing time of the proposed method is minimum than the previous classifier.

Moreover, [35] utilized the Correlation-based Feature Selection technique and used the ALL-IDB2 database. Furthermore, 130 blast lymphocyte images and 130 normal lymphocyte single-cell images are included in the database. Additionally, each image extracted 21 shape features. The research result achieved an accuracy of 92.30% with a set of sixteen features. This result is based on 260 blood microscopic images of lymphocyte. Also, faster detection can be obtained by reducing the number of features makes.

Additionally, [36] utilized a deep convolutional neural network. This research employed the ALL-IDB 2 data set contained 260 images having a single cell. These images consist of 130 images from patients infected by leukemia and 130 normal images. Moreover, these images had a resolution of 257 _ 257 with 24-bit color depth. The results show an average sensitivity of 100%, the specificity of 98.11%, and accuracy of 99.50% for ALL detection. For ALL subtype classification, the experiments performed an average sensitivity of 96.74%, the specificity of 99.03%, and accuracy of 96.06%. Therefore, the proposed method has better results than previous standard methods as there were not image segmentation. Also, [37] applied a Deep Convolutional Network and used the ALL-IDB dataset. 105 are labeled properly out of accessible 108 patches. The proposed method obtained objectively well on patches with a single isolated White Blood Cells (WBC). However, it fails for patches with two or more WBC cells sticking together.

Furthermore, [38] conducted image acquisition, image preprocessing, image segmentation, and feature extraction. This research utilized microscopic blood cell images. Furthermore, the mean and standard deviation of blast cells are smaller than normal blood cells as non-uniform shapes and structure of blast cells. Besides, the elimination of unwanted details of RBCs and blood platelets have been conducted using histogram equalization and contrast enhancement. Additionally, the nucleus of the WBCs cell is highlighted using the Otsus thresholding method. This threshold image of WBCs cell is employed for edge detection. The statistical feature of blast cells helps to differentiate between a

normal cell and blast cell. Moreover, calculation of area and perimeter of blood cell images for the next feature can be conducted using the edge detected images.

Also, [39] applied A fuzzy-based classifier and used a set of 32 blood smears. The research results obtained 93.75% of accuracy. Additionally, the experiment results pointed out that the proposed method is reliable, smaller errors, higher accuracy, less time-consuming, lower cost, as well as efficient. Moreover, [40] employed image processing for examining morphological, textural, and color features from the blood microscopic images. The data set was ALL-IDB. This system has been tested using 108 images of the ALL-IDB1 dataset and detected white blood cells and classified them with the trained Ensemble Classifier. It is based on the 70 features (50s are major) of 260 white blood cells (ALL-IDB2. The system has performed a better accuracy of 93.6% and a sensitivity of 96%.

Besides, [41] utilized a Gray level co-occurrence matrix and random forest-based for ALL detection. The research obtained accuracy results at 96.29% of segmentation, 99.004% of classification and 96 % for cytoplasm and nucleus in each specific part, [42] applied morphology and CD marker for the research. It performed 99.67% accuracy which has better results than other existing methods, [43] employed color and size based objects and machine learning-based for classification method. This method obtained 91.6% precision for the 40× objective and 79.7% for the 20× objective, [44] applied a local pixel information method, and used the ALL-IDB image database. The dataset contained 108 microscopy blood images: 49 abnormal (blast cells) and 59 normal WBCs. The developed segmentation technique identified 630 cells out of 649 cells with an accuracy of 97.07%. Hence, the proposed method has better accuracy than others, also [45] utilized Two Bare-bones Particle Swarm Optimization (BBPSO) algorithms and employed ALL-IDB2 database. The proposed algorithms obtained superior geometric mean performances of 94.94% and 96.25%, respectively, and better results than other metaheuristic search and related methods especially for ALL classification.

It proves that any information has been ignored by two other components as supporting the feature generator. We extract densitometry for all of the channels to overcome the loss information, i.e., red, green, and blue channels, so the features captured can be represented the real object before the classification process. The purpose of this paper is to segment the white blood cell image and classify it as to whether someone is infected with Acute Lymphoblastic Leukemia.

The classification result depends on some factors, i.e., pre-processing, segmentation, and feature extraction. Most of the researchers have ignored important factors, as mentioned before [15, 19, 25], even though an intelligence-based measurement has applied; however, the obtained accuracy cannot achieve maximum.

In this research, the color-based hybrid is employed to achieve the dominant object features, i.e., shape and color based densitometry. The shape features include area, perimeter, and circularity features. Furthermore, the color based densitometry has three elements, for each item has four features, i.e., Energy, Entropy, Correlation, and Variance. Moreover, the feature extraction delivers fifteen elements, i.e., three features are created by the shape model, while the remaining is produced by red, green, and blue color-based densitometry. The shape features are the Area (A), Perimeter (P), and Circularity (C). The red color formed the Energy (E), Entropy (En), Correlation (Co), and Variance (Var). The green color also generates four features, i.e., the Energy (E), Entropy (En), Correlation (Co), and Variance (Var). The blue color also shapes four features as formed by red or green color. The details will be explained in the proposed method section.

Additionally, the research results have been compared to other algorithms, as shown in Table 2. The comparison points out that the proposed model produced higher accuracy than the others, i.e., Naïve Bayesian Classifier [25], Color Correlation [19], Fuzzy-based Leukemia Detection [30], Hausdrof SVM-based Leukemia Detection [46], and Automated Differential [15]. An Automated Differential employed three intelligence measurements, i.e., Learning Vector Quantization (LVQ), Multi-Layer Perceptron (MLP), and Support Vector Machine (SVM). As a result, the proposed method is outperformed than those methods.

The structure of the paper is written as follows: Firstly, the introduction covers the background, aim of the research, literature review, and brief of the proposed method. Secondly, the related work section will capture the methods used to resolve the proposed approach. Thirdly, the proposed model section describes a novelty of the approach to enhance, segment, extract, and classify the object. Fourthly, the experimental results and discussion section presents experimental materials used and obtained as well as performed the analysis of the results of the proposed model. Lastly, the conclusion section explains the experimental and discussion results all manuscripts must be in English. These guidelines

include complete descriptions of the fonts, spacing, and related information for producing your proceedings manuscripts.

2. Related work

Classification tasks of the image always related to segmentation and feature extraction, whereas to segment the object, it is required pre-process to enhance the image. Many methods have been developed to improve the quality of the image, such us Histogram Equalization, Adaptive Histogram Equalization, Contrast Limited Adaptive Histogram Equalization, and also Retinex methods. However, not all methods are suitable implemented to improve the quality of the Acute Lymphoblastic Leukemia Image Database (ALL-IDB). Inappropriateness in applying the pre-processing method will cause mismatch results in segmentation. Histogram equalization is an image enhancement method to distribute the grayscale image so that the image has more even intensity distribution than the original image. Consider the image has grayscale x, whereas n_i states frequencies of ith grayscale (variable of i states an index of p_x and n, where the upper and the lower bound of ith index are 0 and L. the probability of the grayscale to the image size can be stated as follows

$$p_x(i) = \frac{n_i}{n}, \qquad 0 \le i \le L \qquad (1)$$

The parameter of n represents the image size, whereas p_x states the image's histogram. The cumulative distribution function as well-known as *cdf* can be defined as follows

$$cdf_x(i) = \sum_{j=0}^{i} p_x(j) \qquad (2)$$

The value of j reflects an index, where the lower and upper bound are 0 and i as the maximum value of j. The new value of the histogram equalization results in h with v^{th} index can be calculated by round-up of the following equation

$$h(v) = \left\lceil \frac{cdf(v)-1}{H \times W - 1} \times (2^8 - 1) \right\rceil \qquad (3)$$

In Eq. (3), the size of the image influences the histogram equalization results, i.e. the height (H) and weight (W) of image. The cumulative distribution function and the number of bits of an image also give impact to the Eq. (3) result. In this case, the histogram equalization always processes separately for each channel. Therefore, the maximum pixel intensity
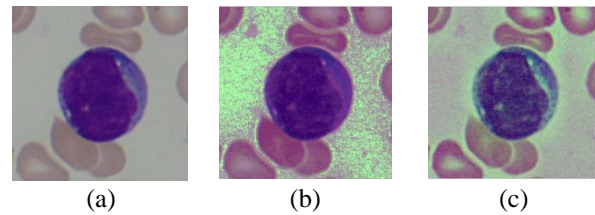

(a)   (b)   (c)

Figure. 1 (a) the original image, (b) histogram equalization result, and (c) contrast limited adaptive histogram equalization result

is 2^8-1 = 256-1=255. If the object processed is a color image, then the Eq. (3) will be operated for each channel, i.e. the red, green, and blue channel. For each channel has 0 for minimum and 255 for maximum intensity.

Histogram equalization does not only produce the better primary object, but also other objects and background has changed clearer as shown in Fig. 1, which has three parts, i.e. the original image, Histogram Equalization result, and Contrast Limited Adaptive Histogram Equalization result (the first, second, last column respectively). The histogram equalization image result delivered a problem when the segmentation process is applied, as around the main object appear a lot of small objects, so that the segmentation process could not produce an appropriate image. A similar result is also provided by using Contrast Limited Adaptive Histogram Equalization. Therefore it is required another method to enhance the image so that the primary object is accessible to be correctly segmented.

In this case, we proposed a model to increase the contrast of the primary and reduce the other objects. This problem is how to determine the best threshold as the separator value between the primary object and background. In this research, the Otsu method is carried out and implemented to the green channel image, and the result is re-combined to the original image. The green channel has been selected for the segmentation because it produces a better contrast image than the red and blue channel (it will discuss more comprehensive on the pre-processing section), while the Otsu method can determine the threshold adaptively by considering the distribution and variation of the image pixels. The image enhancement results show that the desired object has appeared, whereas the others are experiencing a decrease in contrast so that the desired object and background can be easily separated.

3. Proposed model

The primary process of the proposed method is divided into four stages, i.e., Pre-processing,
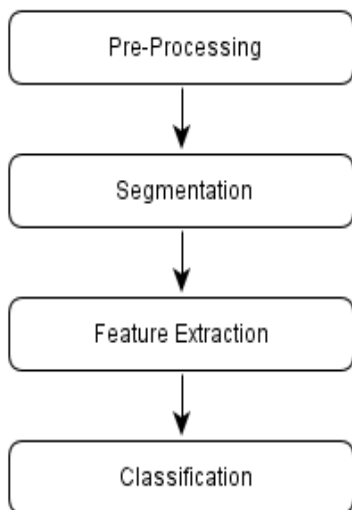
Figure. 2 The stage of the proposed model in classification the acute Lymphoblastic Leukemia
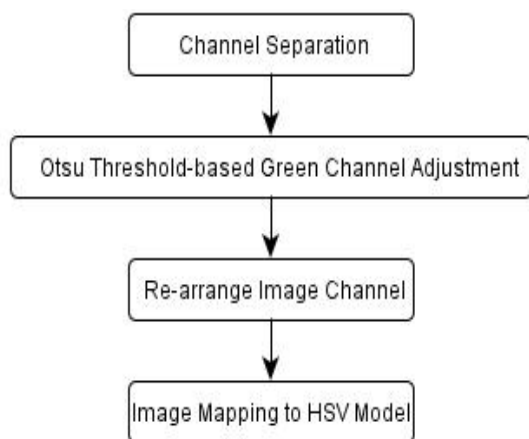

Figure. 3 The proposed of the pre-processing stage

Segmentation, Feature Extraction, and Classification, as displayed in Fig. 2. Pre-Processing is the stage used to enhance the image input before the segmentation step is performed. Segmentation is a process to divide between the object and image background. Feature extraction is a method to obtain the principal characteristics of the object. Classification is a way to identify whether data testing was categorized as Leukemia or not.

3.1 Pre-processing

Pre-processing has a critical role in segmentation and classification. The Segmentation results will be less accurate when the Pre-processing was not correctly conducted. However, the segmentation results also depend on the Pre-processing results.

There are four stages in the pre-processing, i.e., channel separation, Otsu threshold-based green channel adjustment, re-arrange image channel and image mapping to the HSV model. The proposed approach of the Pre-processing stage can be described in Fig. 3. Every step is explained in detail in the following subsection.

3.1.1. Channel separation

In this research, the image inputs use the color image. Suppose $f(x, y, z)$ represents the color image, and $z \in 1$, 2, and 3 represent red ($rf(x, y)$), green ($gf(x, y)$), and blue image channel ($bf(x, y)$). Furthermore, the separation of the color image can be represented in Fig. 4. In this figure, the green channel is clearer than the red and blue channels. Therefore, the green channel will be further processed to obtain a better image, i.e., Otsu threshold-based green channel adjustment. An image sample has $2^{8+8+8} = 2^{24}$ pixel combinations, i.e., 8 bit for red, 8 bit for green, and 8 bit for blue channels. If these channels are separated, then the green channel produces a more contrast image than the other channel, i.e., red and blue channels. We have proven to separate three channels on the image color; evidently, the green channel has produced a sharper image than the others as shown in Fig. 4.

3.1.2. Otsu threshold-based green channel adjustment

The channel separation, as shown in Fig. 4 will deliver an optimum result if the image result is adjusted, where the pixels that tend to be dark are reduced in intensity, while the bright ones are added to the intensity. Adjustment of the green channel has mapped into a new space of the image so that it can produce a better image for both the contrast and quality. In this case, the proposed method is to increase the image contrast based on the threshold found as follows.

$$nf(x, y) = \begin{cases} gf(x, y) + T & gf(x, y) > T \\ gf(x, y) - T & gf(x, y) \leq T \end{cases} \quad (4)$$
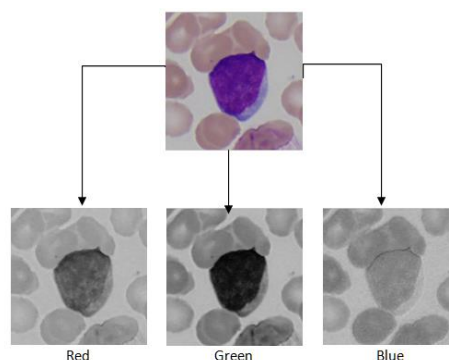

Figure. 4 Source image channel separation (Red, Green, and Blue Channel)

A variable of $nf(x, y)$ represents the mapping result at x and y position, $gf(x, y)$ describes grayscale of the image at x and y position, whereas T denotes the threshold value. The threshold can be found by using the Otsu method. Because the Otsu result produces the range values between 0 until 1, then the results are multiplied with a maximum grayscale image, so that the thresholds have the range between 0 until the maximum grayscale image found. Fig. 5 shows an enhancement result using Eq. (1). It proves that the image enhancement (Fig. 5 (b)) has a more significant contrast than an original image (Fig. 5 (a)).

A threshold is a critical value to separate object and background of the image. The higher the threshold value chosen, the more background obtained. If a threshold value is smaller, then the background acquired cannot expand in the desired direction. In this research, the Otsu method is proposed to achieve the best threshold. The Otsu method utilizes the discriminant value to gain the best value. Consider an image with H as image height and W as image width, whereas the number of intensity and its index is represented by t and k, where the value of $k \in 1...t$. If $f(k)$ and $F(k)$ represent gray level and k^{th} frequency of pixel intensity for background, whereas $\mu_b(s)$ states mean between $k=1$ until the maximum of k, which is s (k states an index of f and F), then $\omega_b(s)$, $\mu_b(s)$, and $\sigma_b{}^2(s)$ describe weight, mean, and variance for the background. These values for the background can be represented as follows.

$$\omega_b(s) = \frac{1}{H \times W}\sum_{k=1}^{S} F\,(k) \qquad (5)$$

$$\mu_b(s) = \frac{\sum_{k=1}^{S}\bigl(f(k)\bullet F(k)\bigr)}{\sum_{k=1}^{t} F(k)} \qquad (6)$$

$$\sigma_b^2(s) = \frac{\sum_{k=1}^{S}\bigl(f(k)-\mu_b(s)\bigr)^2 \bullet F(k)}{\sum_{k=1}^{S} F(k)} \qquad (7)$$

Symbol of "•" showed the inner product of two matrices with the same size. The variable of s is less than or equal to t as shown in Eqs. (5), (6), and (7).

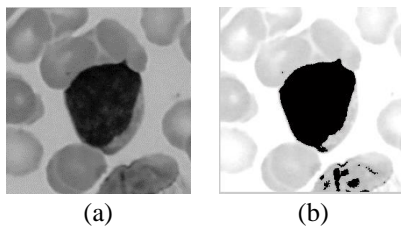<table>
<tr><td>(a)</td><td>(b)</td></tr>
</table>

Figure. 5 Image enhancement: (a) original image and (b) image enhancement result

In Eqs. (8), (9), and (10), the $\omega_f(k)$, $\mu_f(k)$, and $\sigma_f^2(k)$ delineated the weight, mean, and variance for a foreground of the object. If The $f(h)$ and $F(h)$ represent a grayscale and k^{th} frequency of pixel intensity for the foreground, then the weight, mean, and variance can be portrayed as follows

$$\omega_f(k) = \frac{1}{H \times W}\sum_{h=t-k}^{t} F\,(h) \qquad (8)$$

$$\mu_f(k) = \frac{\sum_{h=t-k}^{t}\bigl(f(h)\bullet F(h)\bigr)}{\sum_{h=k}^{t} F(h)} \qquad (9)$$

$$\sigma_f^2(k) = \frac{\sum_{h=t-k}^{t}\bigl(f(h)-\mu_f(t)\bigr)^2 \bullet F(h)}{\sum_{h=t-k}^{t} F(h)} \qquad (10)$$

The symbol of h states an index of f and F, where they have range t-k as lower bound and t as the upper bound. The values of the within and between class variance are represented by $\sigma_W^2(t)$ and $\sigma_B^2(t)$. The within and between class variance can be represented by the following equation

$$\sigma_W^2(t) = \omega_b(t) \times \sigma_b^2(t) + \omega_f(t) \times \sigma_f^2(t) \qquad (11)$$

$$\sigma_B^2(t) = \omega_b(t) \times \omega_f(t) \times \bigl(\mu_b(t) - \mu_f(t)\bigr)^2 \qquad (12)$$

Based on Eqs. (11) and (12), the best threshold can be found from the minimum value of the within class variance or the maximum value of the between class variance.

3.1.3. Re-arrange image channel

The image of the green channel adjustment result is utilized to obtain the best color image to be more easily segmented. To acquire the best color image, an image adjustment result is further used to replace the original green channel so that it can create a new image as in Fig. 6. It showed that the image enhancement result replaces an original green channel. The replacement result displayed the dominant feature is green, whereas the object found is purple. However, the small objects also appear as noises in this image.

3.1.4. Image mapping to HSV model

A further process is to transform from the image result (Fig. 6) to the HSV model, as displayed in Fig. 7. The result appears that the object found has green channel dominant. The mapping results in three components, i.e., *H*, *S*, and *V* represent *hue*, *saturation*, *value* respectively. The *hue* has ranged
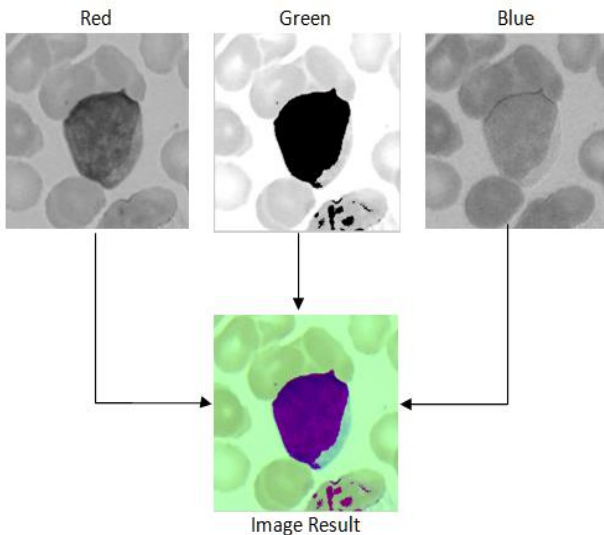
Figure. 6 Replacement green channel of the adjustment result with original red and blue channel

between 0^0 to 360^0, while *the saturation* and *value* have the same range, which is 0^0-100^0. If the *r*, *g*, and *b* state the red, green, and blue channels consecutively, then the mapping from color to *V* model can be stated as follows.

$$V = max(r, g, b) \tag{13}$$

The variable of *V* at x and y position can be described by $V(x, y)$. Therefore the value of $V(x, y)$ can be selected by the maximum values of $r(x, y)$, $g(x, y)$, and $b(x, y)$. In this case, $r(x, y)$, $g(x, y)$, and $b(x, y)$ state the red, green, and blue values at x and y position respectively. Moreover, the saturation at the x and y position is represented by $S(x, y)$. This value can be produced based on the results of $V(x, y)$ as follows

$$S = \begin{cases} 0 & \text{if } V = 0 \\ 1 - \frac{min(r,g,b)}{V} & \text{if } V > 0 \end{cases} \tag{14}$$

For both the values of $V(x, y)$ and $S(x, y)$ as shown in Eqs. (13) and (14) are applied to compute the $V(x, y)$ value as indicated in the following equation.
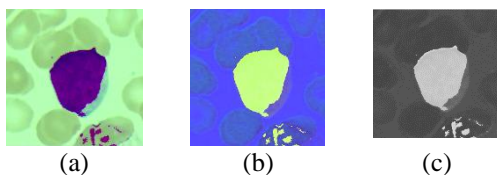

Figure. 7 Transform from image to HSV model: (a) source Image, (b) HSV transformation result, and (c) *Hue* channel
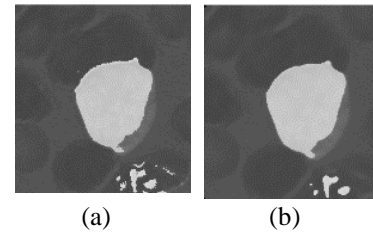

Figure. 8 The median filtering: (a) source image and (b) result image

$$H = \begin{cases} 0, & \text{if } S = 0 \\ \dfrac{60 \times (g - b)}{S \times V} & \text{if } V = r \\ 60 \times \left(2 + \dfrac{b - r}{S \times V} \right) & \text{if } V = g \\ 60 \times \left(4 + \dfrac{r - b}{S \times V} \right) & \text{if } V = b \end{cases} \tag{15}$$

$$H = H + 360 \quad \text{if } H < 0 \tag{16}$$

The mapping from color to the HSV is utilized to obtain the green pixel dominant on the object, as displayed in Fig. 7. It proved that the object found has a green color. Finally, the *Hue* channel is used to further process, as shown in Fig. 7.

3.2 Segmentation

However, the small noises are necessary to merge to another object. Therefore the filter process is required to obtain a better result. In this case, Median Filtering is performed to overcome it. The Median Filtering takes the median value of the neighborhood after the sorting process. The result shows that the small objects can be automatically removed after the filtering process, and it depends on the matrix mask used. The bigger mask matrix size used, the more small objects removed as shown in Fig. 8.

3.2.1. Binary image transformation

The next stage is to transform into a binary image. The problem is how to determine the appropriate value for the threshold; it is applied because each image has a different grayscale distribution. Therefore it is required the method to overcome it. Errors in determining the threshold value will have an impact on the results obtained. In this case, the Otsu method is also applied to achieve the best threshold, as shown in Eqs. (5) until (12). The binary image transformation is shown in Fig. 9.
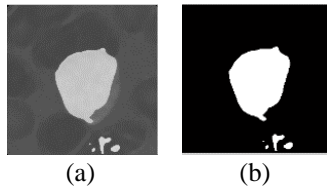
(a)   (b)

Figure. 9 Binary image transformation: (a) source image and (b) binary image result

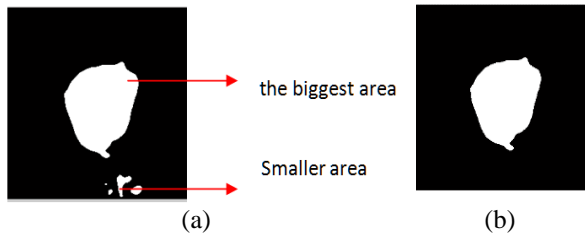(a)   (b)

Figure. 10 Selection the biggest area: (a) object selection process and (b) image results with an object

3.2.2. The biggest object area selection

The binary image transformation still leaves the problem, i.e., there is some noise. Therefore these must be removed from the selection of the maximum area of the image. In this research, the biggest will be applied to a further process. Therefore the binary image result must be selected to determine one of many objects found in the image. The simple algorithm was developed to choose it, i.e., by comparing the object areas in the image. Suppose n objects are in the image, and for an area of the object is represented by using O_k, where k and n state the number of objects and the area index. The biggest area is chosen as an object candidate, as shown in the following model.

$$M = max(O_1, O_2, O_3, O_4, \cdots, O_n) \qquad (17)$$

Fig. 8 showed an image with five objects, and it means that $n=5$, $k=1, 2, 3, 4, 5$. Five objects will be selected one object with the biggest area, and the object result is displayed in Fig. 10 (b).

3.2.3. Re-arrangement of the source and the negative of the binary image result transformation

Before the feature extraction process, the image segmentation result must be returned to become the color image. It is necessary to create the negative of the binary image to obtain the color image result. The simple mathematics can be modeled to produce the negative of the binary image as follows
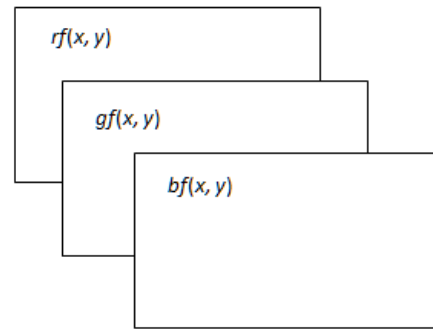
$$nB(x, y) = 255 - B(x, y) \qquad (18)$$

Figure. 11 Re-arrangement of the source and the negative of the binary image

$nB(x, y)$ and $B(x, y)$ state the negative of a binary image and the biggest. The value of $nB(x, y)$ is used as a deduction against the original channel, i.e., red, green, and blue channel as follows

$$rf(x, y) = f(x, y, 1) - nB(x, y) \qquad (19)$$

$$gf(x, y) = f(x, y, 2) - nB(x, y) \qquad (20)$$

$$bf(x, y) = f(x, y, 3) - nB(x, y) \qquad (21)$$

The values of the $rf(x, y)$, $gf(x, y)$, and $bf(x, y)$ display the new value of the red, green, and blue based on the Eq. (18). The results of Eqs. (19), (20), and (21) are further re-arranged in the color format so that it can produce the color image, as shown in Fig. 11. It shows that the color image result can be re-arranged based on the results of Eqs. (19), (20), and (21).

Fig. 12 explains the formation of the color image using the Eqs. (19), (20), and (21). Each channel is subtracted with a negative binary image, and the results are combined to obtain a single color image.
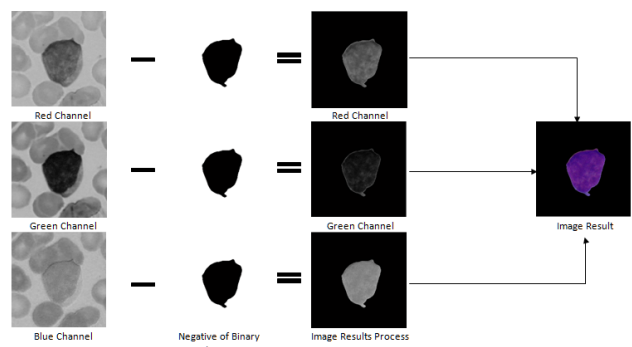

Figure. 12 Re-arrangement of the subtraction results of the source image and the negative of binary image result transformation

3.3 Color-based hybrid feature extraction modeling

An essential process after pre-processing and segmentation is feature extraction. This process also determines how well the classification results. In this research, the color-based hybrid is utilized to obtain the dominant object features, i.e., shape and color based densitometry. In general, densitometry is applied to the grayscale image, therefore, if a color image is characterized based on the grayscale image, then some features will disappear, and the object will also lose the original characteristics. In this research, the color based hybrid is proposed to extract the object features. In this case, a phrase of "color based hybrid" has the meaning of combining the shape and color based densitometry features. The shape features involve Area, Perimeter, and Circularity features, whereas the color based densitometry has three elements, for each item has four features, i.e., Energy, Entropy, Correlation, and Variance.

3.3.1. The shape features

The shape features consist of three characteristics, i.e., Area (A), Perimeter (P), and Circularity features (C). An object area can be computed by reading the image segmentation results. Suppose, binary image $B(x, y)$ states intensity on the x and y position if $B(x, y) \leftarrow 1$, the value of area increases 1. The image intensity is read from the top-left until the bottom-right. If $B(x, y) \leftarrow 1$ and it has at least a neighbor with the zero intensity, then the perimeter increases 1. It is to compute the object perimeter. The object circularity can be obtained by using the following model

$$C = \frac{4 \times \pi \times A}{P^2} \tag{22}$$

The variable of π represent and phi value (it is 3.14)

3.3.2. The color-based densitometry features

Another feature is the densitometry. In general, the researcher only applied a green channel to extract the object features. Therefore the object has the essential characteristic as the features. In this research, three channels of the object will be extracted to obtain the dominant features, i.e., Densitometry for red, green, and blue. It measures intensity uniformity, the complexity of the size, interconnection between a pixel to nearest neighbors, and pixel variance of the image. The intensity uniformity measurement can be represented by using the "Energy (E)" as follows:

$$Energy = \sum_{j=0}^{H-1} \sum_{k=0}^{W-1} (g(j,k))^2 \tag{23}$$

In this case, H, W, j, k, and $g(j, k)$ represent the height, wide, index of the height, index of the width of the image, and new space of the image respectively. The intensity uniformity can be calculated by multiplication of the source image and log of a source image. For each position must be calculated the value of "Entropy (En)" and the results must be added, so that it produces the single value as follows:

$$Entropy = -\sum_{j=0}^{H-1} \sum_{k=0}^{W-1} g(j,k) \times log(g(j,k)) \tag{24}$$

A relation of the pixel and nearest neighbor can be represented by using the "Correlation (Co)" function as follows

$$Correlation = \sum_{j=0}^{H-1} \sum_{k=0}^{W-1} \frac{(j-\mu_x) \times (k-\mu_y)}{\sigma_x \times \sigma_y} \tag{25}$$

The variable of the μ_x, μ_y, σ_x, and σ_y depict the average of the x, y, variance of the x, and y respectively. The last feature of the densitometry is the "Variance (Var)", where the mean image value μ must be calculated before the variance σ, as shown in the following equation.

$$Variance = \sum_{j=0}^{H-1} \sum_{k=0}^{W-1} (j - \mu) \times g(x, y) \tag{26}$$

In this research, the Eqs. (23), (24), (25), and (26) are applied on the different channels, i.e., red, green, and blue channels. Therefore the color based densitometry has twelve features. Some features used are fifteen, i.e., the 1st, 2nd, and the 3rd are features of an Area, Perimeter, and Circularity, respectively, the 4th, 5th, 6th, and 7th state features of the Energy, Entropy, Correlation, and Variance for the red channel. For the green channel is represented by the 8th, 9th, 10th, and 11th features, while for the blue channel is 12th, 13th, 14th, and 15th features as shown in Fig. 13

The feature extraction produces fifteen elements, i.e., three features are formed by the shape model,

The Color based Hybrid Modelling														
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	P	C	E	En	Co	Var	E	En	Co	Var	E	En	Co	Var

The shape features   The red channel features   The Green channel features   The blue channel features

Figure. 13 Color-based hybrid modeling (The shape and the color based densitometry)

while the remaining is generated by red, green, and blue color-based densitometry. The shape features are the *Area* (*A*), *Perimeter* (*P*), *and Circularity* (*C*), as shown on the 1st, 2nd, and 3rd columns. The red color fashioned the *Energy* (*E*), *Entropy* (*En*), *Correlation* (*Co*), and *Variance* (*Var*) as described on the 4th, 5th, 6th, and 7th columns. The green color also forms four features, i.e., the *Energy* (*E*), *Entropy* (*En*), *Correlation* (*Co*), and *Variance* (*Var*), as exhibited on the 8th, 9th, 10th, and 11th columns. The blue color also produces four features as generated by red or green color, and the results are 12th, 13th, 14th, and 15th columns.

3.4 Classification

The image sample consists of the training and the testing sets. The feature extraction results will be compared and calculated the accuracy. The measurements are performed by comparing the training and the testing sets.

In the classification stage, there are two sub-stages, i.e., Similarity Measurements and Accuracy Calculation. On the Similarity Measurements, the training features produced will be employed as a reference to the testing process. The samples are performed by using the same process. In this research, two methods are applied to classify image testing, i.e., Manhattan and Euclidean Distance, where *m* represents the number of features used. A sequence of the feature extraction has not to default rule, as long as it is consistent. The Manhattan and Euclidean Distance can be easily described as follows.

$$D_M = \sum_{L=1}^{m} |Tr_L - Ts_L| \tag{27}$$

$$D_E = \sqrt{\sum_{L=1}^{m} (Tr_L - Ts_L)^2} \tag{28}$$

The variable of D_M, D_E, Tr_L, Ts_L, *L*, and *M* represent the similarity result using Euclidian distance, the similarity result using Manhattan, feature extraction of the training set, feature extraction of the testing set, index of the feature extraction, and number of feature extractions respectively.

The calculation results of the Eqs. (27) and (28) are moreover applied to calculate the accuracy of the measurement results. In this case, Accuracy, (*ACC*), True Positive Rate or Sensitivity (*TPR*), and False Negative Rate or Specificity (*FNR*) must be calculated as follows.

$$TPR = \frac{TP}{TP+FN} \tag{29}$$

$$TNR = \frac{TN}{TN+FP} \tag{30}$$

$$ACC = \frac{TPR+TNR}{2} \tag{31}$$

True Positive (*TP*) states positive Acute Lymphoblastic Leukemia correctly identified as an abnormal image. False Positive (*FP*) represents healthy Acute Lymphoblastic Leukemia identified as an abnormal cell. True Negative (*TN*) states a normal cell correctly identified as a normal cell, whereas False Negative (*FN*) stated positive Acute Lymphoblastic Leukemia cell identified as the normal cell.

4. Experimental results and discussion

In this research, the Acute Lymphoblastic Leukemia Image Database for Image Processing (ALL-IDB) is utilized to evaluate the proposed model. ALL-IDB used is type-2 or well known as ALL-IDB2. It consists of 260 images, where 130 images are classified as Acute Lymphoblastic Leukemia, and the remaining are signed as a normal white blood cell. These images are designed to classify Acute Lymphoblastic Leukemia. The following is a sample of Healthy (Fig. 14) and Acute Lymphoblastic Leukemia Cells (Fig. 15).

Moreover, the results are applied to calculate Accuracy, Sensitivity, and Specificity. Fig. 16 displays the classification accuracy results using Euclidian Distance.
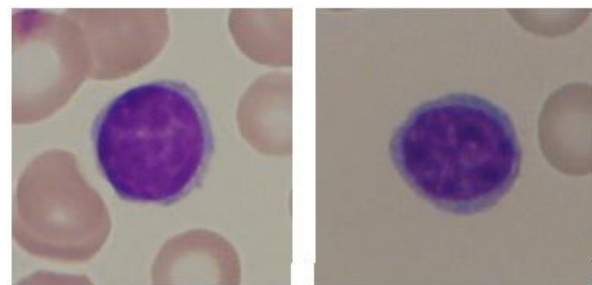

Figure. 14 Healthy cell of acute Lymphoblastic Leukemia image database – ALL-IDB
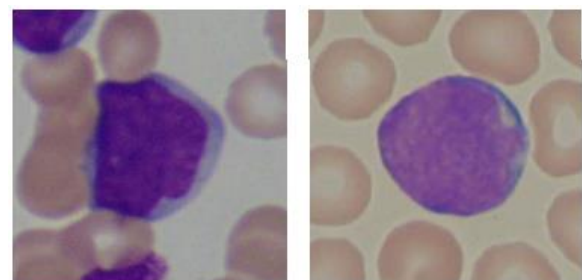

Figure. 15 Acute Lymphoblastic Leukemia cell of ALL-IDB

Table 1. Experiment scenario used

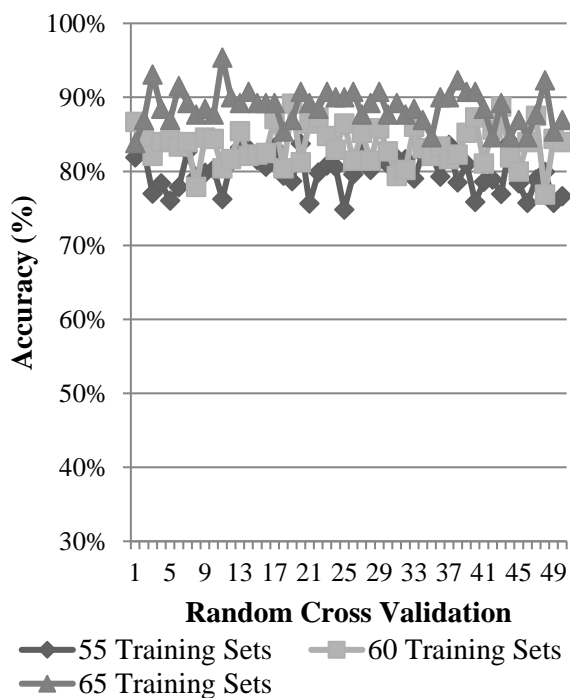| Scenario	Training Sets		Testing Sets	
	ALL	Non ALL	ALL	Non ALL
1st	55	55	75	75
2nd	60	60	70	70
3rd	65	65	65	65

Figure. 16 Accuracy of the proposed model using Euclidian distance

Usage of the training sets had to impact the classification results because the results showed that the usage of 65 training sets had produced higher accuracy than 60 and 55 training sets. For each scenario, experiments are conducted 50 times with the different indexes on the training sets. The low fluctuation of the accuracy has been produced on the 1st, 2nd, and the 3rd scenario, i.e., 0.24, 0.26, and 0.23 standard deviations, while the average accuracy produced is 79.85%, 83.54%, 88.72% for the 1st, 2nd, and 3rd respectively.

Fig. 16 also displayed that the maximum accuracy obtained is 83.97% for the 1st, 89.17% for the 2nd, and 95.38% for the last scenario. Increasing the training sets used has provided an increase in average and maximum classification accuracy obtained. It proves that the proposed model is suitable to be used as a classification of the Acute Lymphoblastic Leukemia.

A similar process has been performed by using the Manhattan method, as shown in Fig. 17. The
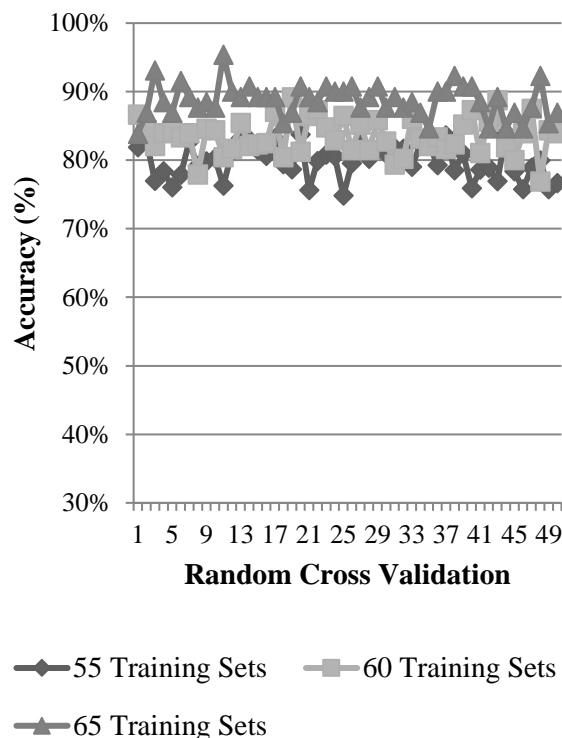

Figure. 17 The accuracy of the proposed model using Manhattan

experimental results have produced the average accuracy, i.e., 79.37% for the 1st, 83.03% for the 2nd, and 88.60% for the 3rd scenario, while the maximum accuracy obtained is 83.97%, 89.17%, 95.38% for the 1st, 2nd, and 3rd respectively. The significant increase has been displayed for both the average and the maximum accuracy. The similar results of the standard deviation have also been shown, i.e., 0.27, 0.22, and 0.23 standard deviations for the 1st, 2nd, and 3rd scenario. It also showed that classification accuracy depends on the training sets used.

Similar results have been produced on the Euclidian Distance and Manhattan for both the average and maximum accuracy. A little difference occurred, where Manhattan outperformed to Euclidian Distance on the 1st scenario, but on the 2nd scenario, the Euclidian Distance produced higher accuracy than the Manhattan, as shown in Fig. 18. Similar results also occur at the maximum accuracy, where Euclidian Distance outperformed to Manhattan when the 2nd and 3rd scenarios are performed, but it did not happen in the 1st scenario. Precisely, the Manhattan exceeded to Euclidian Distance on the 1st scenario, as displayed in Fig. 19.

The experimental results showed that the highest classification accuracy occurred in the last scenario when Euclidian Distance is applied as similarity measurements, i.e., 95.38%. The experimental results

of the proposed model have also been evaluated to be compared to other algorithms, as shown in Table 2. It demonstrated that the proposed model had obtained higher accuracy than the others, i.e., Naïve Bayesian Classifier [25], Color Correlation [19], Fuzzy-based Leukemia Detection [30], Hausdrof SVM-based Leukemia Detection [46], and Automated Differential [15]. An Automated Differential utilized three intelligence measurements, i.e., Learning Vector Quantization (LVQ), Multi-Layer Perceptron (MLP), and Support Vector Machine (SVM). Three method results are not better than the proposed model.

Table 2 showed that the highest accuracy for an intelligence-based approach is SVM, i.e., 91.3%, whereas an intelligence-based method others have produced 83.33% and 89.74%, i.e., LVQ and MLP. It showed that the experimental results of the proposed model are also influenced by three important factors, i.e., Pre-processing, segmentation, and feature extraction.

Another approach has been produced by other researchers, i.e., the statistical-based model (Naïve Bayesian Classifier) [30] only contributed 75% accuracy. It proved that the proposed model had outperformed the others, for both intelligence and statistical-based approach.

Similarity measurements also have a significant role in producing accuracy. However, the features measured are also more important for the classification process. In the feature extraction, segmentation process contributed to producing important features.
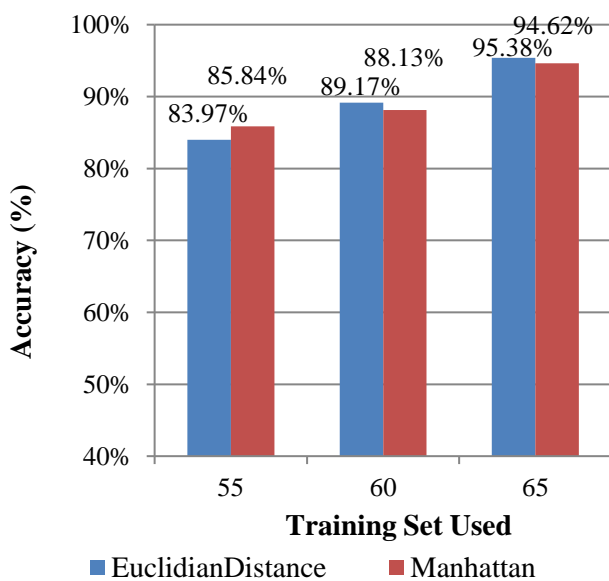

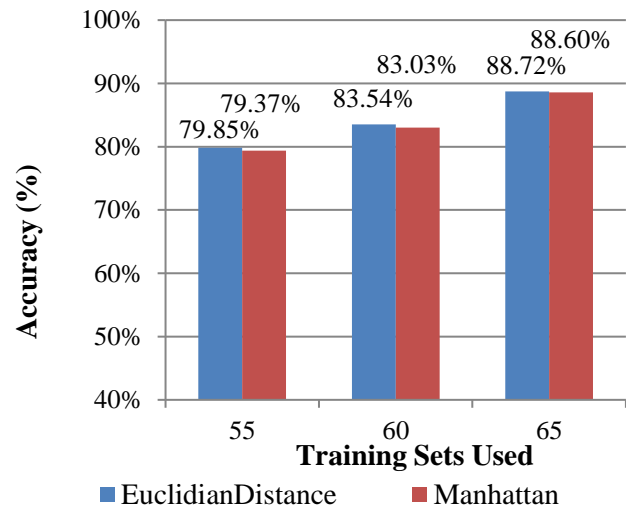Figure. 18 The maximum accuracy of the proposed model using Euclidian distance

EuclidianDistance   Manhattan

Figure. 19 The maximum accuracy of the proposed model using Manhattan

Table 2. Comparison of the proposed model and the others

No	Method	Accuracy (%)
1	Proposed Model	95.38
2	Naïve Bayesian [25]	75,00
3	Color Correlation [19]	92.00
4	Fuzzy-based Leukemia Detection [30]	93.00
5	Hausdrof SVM-based Leukemia Detection [46]	95.00
6	Automated Differential [15]	
	LVQ	83.33
	MLP	89.74
	SVM	91.03

5. Conclusion

A sequential process to enhance the image, separate the desired object, extract the features, and measure the similarity has been conducted by using the proposed model. The simple algorithm to enhance the image has shown satisfactory results. A new approach to separate the object has also displayed the desired object. The segmentation results have been extracted by using the shape and the color based densitometry, while the densitometry feature extraction was conducted by using multi-channel, i.e., red, green, and blue. The results showed that the maximum accuracy obtained is 95.38%, and this result has also outperformed to the others, i.e., Naïve Bayesian at 75%, Color Correlation at 92%, Fuzzy-based Leukemia at 93%, Hausdrof SVM-based Leukemia Detection at 95%, LVQ at 83.33%, MLP at 89.74%, and SVM at 91.03%. This research makes a contribution by providing a new approach to

enhance the image, segment, extract, and measure the similarity, particularly on the Acute Lymphoblastic Leukemia Images. This novel approach has been proven that has better accuracy than others as explained above.

In some cases, the leukemia image still leaves noise before the feature is extracted, therefore we will improve the segmentation process to reduce the noise in the future work. Also we enhance the feature extraction to produce tha main feature that can distinguish postive and negative leukemia.

Conflicts of Interest

We would like to declare no conflicts of interest.

Author Contributions

The authors contributions are as follows: "Conceptualization, Arif Muntasa; methodology, Arif Muntasa; software, Arif Muntasa; validation, Arif Muntasa; formal analysis, Arif Muntasa; investigation, Arif Muntasa; literature reviews, Arif Muntasa and Muhammad Yusuf;resources, Arif Muntasa; data curation, Arif Muntasa; writing— original draft preparation, Arif Muntasa; writing— review and editing, Arif Muntasa and Muhammad Yusuf; revising, Arif Muntasa and Muhammad Yusuf; corresponding, Muhammad Yusuf; Proofreading, Muhammad Yusuf;visualization, Arif Muntasa; supervision, Arif Muntasa; project administration, Arif Muntasa and Muhammad Yusuf; funding acquisition, Arif Muntasa and Muhammad Yusuf".

References

[1] G. Läthén, J. Jonasson, and M. Borga, "Blood vessel segmentation using multi-scale quadrature filtering", *Pattern Recognit. Lett.*, Vol. 31, No. 8, pp. 762–767, 2010.

[2] A. Muntasa, I. A. Sirajudin, and M. K. Sophan, "Matrix mask overlapping and convolution eight directions for blood vessel segmentation on fundus retinal image", *Telkomnika (Telecommunication Comput. Electron. Control.*, Vol. 12, No. 3, pp. 631–638, 2014.

[3] M. U. Akram, A. Tariq, and S. A. Khan, "Retinal image blood vessel segmentation", In: *Proc. of International Conference on Information and Communication Technologies*, pp. 1–5, 2009.

[4] M. Park, J. S. Jin, and S. Luo, "Locating the optic disc in retinal images", In: *Proc. of the International Conference on Computer Graphics, Imaging and Visualisation*, pp. 141–145, 2006.

[5] C. A. Perez, D. A. Schulz, C. M. Aravena, C. I. Perez, and V. Juan, "A new method for online retinal optic-disc detection based on cascade classifiers", In: *Proc. of IEEE International Conference on Systems, Man, and Cybernetics*, pp. 4300–4304, 2013.

[6] J. Xu, O. Chutatape, and P. Chew, "Automated optic disk boundary detection by modified active contour mode", *IEEE Trans. Biomed. Eng.*, Vol. 54, No. 3, pp. 473– 482, 2007.

[7] R. J. Qureshi, L. Kovacs, B. Harangi, B. Nagy, T. Peto, and A. Hajdu, "Combining algorithms for automatic detection of optic disc and macula in fundus images", *Comput. Vis. Image Underst.*, Vol. 116, No. 1, pp. 138–145, 2012.

[8] J. Cavallereno and L. M. Aiello, "Emerging trends in ocular telemedicine: the diabetic retinopathy model", *J. Telemed Telecare*, Vol. 11, No. 4, pp. 163–166, 2005.

[9] R. M. Davis, S. Fowler, K. Bellis, J. Pocki, V. A. Pakalnis, and A Woldorf, "Telemedicine improves eye examination rates in individuals with diabetes", *Diabetes Care*, Vol. 26, No. 8, pp. 165–173, 2003.

[10] A. Rejintal and N. Aswini, "Image Processing Based Leukemia Cancer Cell Detection", In: *Proc. of IEEE International Conference On Recent Trends In Electronics Information Communication Technology*, pp. 471–474, 2016.

[11] J. Su, B. Yoon, and E. R. Dougherty, "Accurate and reliable cancer classification based on probabilistic inference of pathway activity", *PLoS One*, Vol. 12, No. 4, p. e8161, 2009.

[12] M. Wang, X. Zhou, F. Li, J. Huckins, R. King, and S. Wong, "Novel Cell Segmentation and Online Learning Algorithms for Cell Phase Identification in Automated Time-lapse Microscopy", In: *Proc. of the 4th IEEE International Symposium Biomedical Imaging (ISBI)*, pp. 65–68, 2007.

[13] A. Nasir, M. Mashor, and H. Rosline, "Unsupervised Colour Segmentation of White Blood Cell for Acute Leukaemia Images", In: *Proc. of IEEE International Conference on Imaging Systems and Techniques (IST)*, pp. 142–14, 2011.

[14] W. Shitong and W. Min, "A new detection algorithm based on fuzzy cellular neural

networks for white blood cell detection", *IEEE Trans. Inf. Technol. Biomed.*, Vol. 10, No. 1, pp. 5–10, 2006.

[15] G. Ongun, U. Halici, K. Leblebicioglu, V. Atalay, M. Beksac, and S. Beksac, "Feature extraction and classification of blood cells for an automated differential blood count system", In: *Proc. of IEEE International Joint Conference in Neural Networks (IJCNN)*, pp. 2461–2466, 2001.

[16] B. Swolin, P. Simonsson, S. Backman, I. Löfqvist, I. Bredin, and M. Johnsson, "Differential counting of blood leukocytes using automated microscopy and a decision support system based on artificial neural networks – evaluation of DiffMasterTM Octavia", *Int. J. Lab. Hematol.*, Vol. 25, No. 3, pp. 139–147, 2003.

[17] S. Agaian, M. Madhukar, and A. Chronopoulos, "Automated Screening System for Acute Myelogenous Leukemia Detection in Blood Microscopic Images", *IEEE Syst. J.*, Vol. 8, No. 3, pp. 995–1004, 2014.

[18] S. Mohapatra, D. Patra, and S. Satpathi, "Automated Cell Nucleus Segmentation and Acute Leukemia Detection in Blood Microscopic Images", In: *Proc. of International Conference on Systems in Medicine and Biology (ICSMB)*, pp. 49–54, 2010.

[19] F. Scotti, "Robust Segmentation and Measurements Techniques of White Cells in Blood Microscope Images", In: *Proc. of the IEEE Instrumentation and Measurement Technology Conference*, pp. 43–48, 2006.

[20] L. Yang, P. Meer, and D. J. Foran, "Unsupervised segmentation based on robust estimation and color active contour models", *IEEE Trans. Inf. Technol. Biomed.*, Vol. 9, No. 3, pp. 475 –486, 2005.

[21] K. S. Kim, P. K. Kim, J. J. Song, and Y. C. Park, "Analyzing Blood Cell Image do Distinguish its abnormalities", In: *Proc. of ACM International Conference on Multimedia*, pp. 395–397, 2002.

[22] T. Markiewicz, S. Osowski, B. Marianska, and L. Moszczynski, "Automatic Recognition of the Blood Cells of Myelogenous Leukemia using SVM", *IEEE International Joint Conference in Neural Networks (IJCNN)*, pp. 2496–501, 2005.

[23] V. Piuri and F. Scotti, "Morphological Classification of Blood Leucocytes by Microscope", In: *Proc. of IEEE International Conference on Computational Intelligence for Measurement Systems and Applications*, pp. 103–108, 2004.

[24] F. Scotti, "Automatic morphological analysis for acute leukemia identification in peripheral blood microscope images", In: *Proc. of IEEE International Conference on Computational Intelligence for Measurement Systems and Applications*, pp. 96–101, 2005.

[25] S. Sriram and K. Bommannaraja, "Naïve Bayesian classifier for Acute Lymphocytic Leukemia detection", *ARPN J. Eng. Appl. Sci.*, Vol. 10, No. 16, pp. 6888–6892, 2015.

[26] L. Putzu and C. Di Ruberto, "White Blood Cells Identification and Classification from Leukemic Blood Image", In: *Proc. of the IWBBIO International Work*, No. January, pp. 18–20, 2013.

[27] R. Mohammed, O. Nomir, and I. Khalifa, "Segmentation of Acute Lymphoblastic Leukemia Using CY Color Space", *Int. J. Adv. Comput. Sci. Appl.*, Vol. 5, No. 11, pp. 99–101, 2014.

[28] M. D. Joshi, A. . Karode, and S. Suralkar, "Detection of Acute Leukemia Using White Blood Cells Segmentation Based on Blood Samples", *Int. J. Electron. Commun. Eng. Technol.*, Vol. 4, No. 3, pp. 148–153, 2013.

[29] R. Bhattacharjee and L. M. Saini, "Detection of Acute Lymphoblastic Leukemia using Watershed Transformation Technique", In: *Proc. of International Conference on Signal Processing, Computing and Control*, pp. 383–386, 2015.

[30] S. Mohapatra, S. S. Samanta, D. Patra, and S. Satpathi, "Fuzzy Based Blood Image Segmentation for Automated Leukemia Detection", In: *Proc. of International Conference on Devices and Communications (ICDeCom)*, pp. 1–5, 2011.

[31] N. Chatap and S. Shibu, "Analysis of blood samples for counting leukemia cells using Support vector machine and nearest neighbour", *IOSR J. Comput. Eng.*, Vol. 16, No. 5, pp. 79–87, 2014.

[32] H. Vaghela, H. Modi, M. Pandya, and M. Potdar, "Leukemia Detection using Digital Image Processing Techniques", *Int. J. Appl. Inf. Syst.*, Vol. 10, No. 1, pp. 43–51, 2015.

[33] M. D. Joshi, A. H. Karode, and S. R. Suralkar, "White Blood Cells Segmentation and Classification to Detect Acute Leukemia", *Int. J. Emerg. Trends Technol. Comput. Sci.*, Vol. 2, No. 3, pp. 147–151, 2013.

[34] B. K. Das and H. S. Dutta, "Infection level identification for leukemia detection using optimized Support Vector Neural Network", *Imaging Sci. J.*, Vol. 67, No. 8, pp. 417–433, 2019.

[35] V. Singhal and P. Singh, "Correlation based feature selection for diagnosis of acute lymphoblastic leukemia", In: *ACM Int. Conf. Proceeding Ser.*, Vol. 10-13-Augu, pp. 5–9, 2015.

[36] S. Shafique and S. Tehsin, "Acute lymphoblastic leukemia detection and classification of its subtypes using pretrained deep convolutional neural networks", *Technol. Cancer Res. Treat.*, Vol. 17, pp. 1–7, 2018.

[37] A. Ghosh, S. Singh, and D. Sheet, "Simultaneous localization and classification of acute lymphoblastic leukemic cells in peripheral blood smears using a deep convolutional network with average pooling layer", In: Proc. of *2017 IEEE Int. Conf. Ind. Inf. Syst.*, Vol. 2018-Janua, pp. 1–6, 2018.

[38] S. Khobragade, D. D. Mor, and C. Y. Patil, "Detection of leukemia in microscopic white blood cell images", In: *Proc. - IEEE Int. Conf. Inf. Process. ICIP 2015*, pp. 435–440, 2016.

[39] M. A. Khosrosereshki and M. B. Menhaj, "A fuzzy based classifier for diagnosis of acute lymphoblastic leukemia using blood smear image processing", In: *Proc. of the 5th Iran. Jt. Congr. Fuzzy Intell. Syst. - 16th Conf. Fuzzy Syst. 14th Conf. Intell. Syst. CFIS 2017*, pp. 13–18, 2017.

[40] A. Rahman and M. M. Hasan, "Automatic Detection of White Blood Cells from Microscopic Images for Malignancy Classification of Acute Lymphoblastic Leukemia", In: *Proc. of the 2018 Int. Conf. Innov. Eng. Technol.*, pp. 1–6, 2019.

[41] S. Mishra, B. Majhi, P. K. Sa, and L. Sharma, "Gray level co-occurrence matrix and random forest based acute lymphoblastic leukemia detection", *Biomed. Signal Process. Control*, Vol. 33, pp. 272–280, 2017.

[42] 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.

[43] F. Uslu, K. Icoz, K. Tasdemir, and B. Yilmaz, "Automated quantification of immunomagnetic beads and leukemia cells from optical microscope images", *Biomed. Signal Process. Control*, Vol. 49, pp. 473–482, 2019.

[44] 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.

[45] W. Srisukkham, L. Zhang, S. C. Neoh, S. Todryk, and C. P. Lim, "Intelligent leukaemia diagnosis with bare-bones PSO based feature optimization", *Appl. Soft Comput.*, Vol. 56, pp. 405–419, 2017.

[46] S. Mohapatra and D. Patra, "Automated Leukemia Detection using Hausdorff Dimension in Blood Microscopic Images", In: *Proc. of Interact*, pp. 64–68, 2010.