

RESEARCH PAPER

Automated Thalassemia cell image segmentation using hybrid Fuzzy C-Means and K-Means

Nabeel J. Ali¹, Sardar P Yaba²

^{1,2} Department of Physics, College of Education, Salahaddin University-Erbil, 44001, Erbil, IRAQ

ABSTRACT:

Thalassemia is a form of hereditary disease. It's one of the world's most common illnesses. The morphology of red blood cells is most affected by this disorder. This research proposes a new method of automatically segmenting red blood cells from microscopic blood smear images. The study suggests a novel combination of image processing techniques and extensive preprocessing to achieve superior segmentation performance. In this study, the eleven designated color spaces, with six filters and three contrasts enhancing, Fuzzy c-means, and K-means segmentation, have been studied using five evaluation parameters. The evaluation is based on the ground truth image. The Photoshop program performs novel ground truth techniques for multi-object sense (RBC cells). The optimization of the image processing stages was obtained through local image datasets (258 images) from seven thalassemia patients in the Erbil – thalassemia center and five samples of normal blood cells in Children Raparin Teaching Hospital. The image was captured with light intensities (low, medium, high) and with /without a yellow filter in Biophysics Research lab /Education College / Salahaddin University –Erbil. This study found that the best light intensity for image slide capture utilizing a microscope was medium without using a yellow filter with an accuracy of 0.91 ± 0.14 and a performance of 95.34%.

KEYWORDS: Thalassemia, color space conversion, segmentation, Fuzzy c-means, K-means, hybrid techniques.

DOI: <http://dx.doi.org/10.21271/ZJPAS.35.4.03>

ZJPAS (2023) , 35(4);22-33 .

1. INTRODUCTION:

Thalassemia comprises two Greek words: Thalassa, which means sea, and Hema, which implies blood (Sadiq et al., 2021). And it is a form of hereditary genetic disease. Thalassemia is one of the world's most common illnesses (Tyas et al., 2017). and is predominantly present among Southern Asian, African individuals, Italian, Greek, and Middle Eastern (Purwar et al., 2021). Thalassemia is a morphological disorder of the red blood cells (RBCs). It is a disease of the RBC morphology (Marzuki et al., 2017). RBC's primary purpose is to provide oxygen (O₂) to bodily tissues via blood flow through the system (Lavanya and Sushritha, 2017).

Any change in the chains might result in low hemoglobin quantity and impair red blood cell formation (RBC). As a result, their primary role as an oxygen transporter to the entire body may be harmed. The worst situations occur when a particular body region does not receive enough oxygen. The organs may be damaged or malfunction as a result (Rashid et al., 2015) β -Thalassemia is caused by mutations in genes that control hemoglobin production. β -thalassemia is one of the most frequent kinds of thalassemia. Individuals with β -thalassemia minor are healthy and are sometimes misdiagnosed with anemia due to a lack of iron (Çil et al., 2020).

Because RBC morphologies frequently vary owing to changed membrane lipid composition, iron insufficiency, or metabolic problems, information on RBC morphology is

* Corresponding Author:

Nabeel Jalal Ali

E-mail: nabeel.ali@su.edu.krd

Article History:

Received: 04/11/2022

Accepted: 01/01/2023

Published: 30/08 /2023

vital for diagnosing blood diseases. The blood smear test, which evaluates the morphology of stained RBCs under a light microscope, is currently used to diagnose various blood-related conditions. The blood film is generally stained to provide appropriate picture contrast for good microscopic analysis (Lin et al., 2020). Pathologists use a light microscope to detect erythrocyte abnormalities in blood smear pictures as part of their daily practice at the hospital. This subjective evaluation procedure results in laborious, time-consuming, and error-prone work (Ahmad et al., 2018). While CBCs can now be automated, technical breakthroughs like flow cytometry are committed to accurately counting and identifying RBCs at a high rate, but they are also difficult to operate. The traditional blood smear to the modern usage of computer-aided machinery is some other ways to analyze RBC morphology. Image processing has aided in the detection and analysis of numerous scientific projects (Pellegrino et al., 2021). The automated image-processing technique consists of five fundamental processes: image acquisition, image processing, image segmentation, image post-processing, and image analysis (Mohamed and Far, 2012). As a result, compared to normal blood cells, Thalassemia blood cells have a smaller shape and form a characteristic formed (Setsirichok et al., 2012). This study aims to segment techniques that could be applied to the proposed. To produce fully automated segmented RBC of thalassemia images, the current study will employ the potential usage of a mix of global contrast approach and color image segmentation utilizing distinct color spaces. Specifically, which is an important reason to diagnose thalassemia. Doctors, hematologists, pathologists, and medical technologists can use the findings to make a preliminary diagnosis of Thalassemia. Through computer-aided automation, it may be possible to improve traditional ways of evaluating RBCs.

2. RESEARCH METHOD

2.1. The Samples

The proposed system was evaluated on 258 microscopic blood smear images. These images have had a resolution of 3488×2616 . Also, local datasets were composed of 258 images, the red blood cell images derived from

seven thalassemia patients obtained from the Erbil – thalassemia center, and five samples of normal blood cells obtained from Children Raparin Teaching Hospital. The 43 fields were taken from several slides with six physical conditions (intensity and filter). The conditions were: low-intensity without filter (LIWF), medium intensity without filter (MIWF), high-intensity without filter (HIWF), low-intensity yellow filter (LIYF), medium-intensity yellow filter (MIYF), high-intensity yellow filter (HIYF). So, the total image becomes 258 ($=43 \times 6$), which consists of 150 ($=25 \times 6$) thalassemia and 108 ($=18 \times 6$) healthy fields shown in Fig1. Utilization of Leishman staining for the slides in the local databases after preparing the slide of blood cells. Then, image Acquisition was performed using a digital light VanGrand microscopy supported by a digital camera (9 MP). The method for showing microscopic blood smear images involves digitizing the image from an optical image captured with a 100-times (100X) objective, corresponding to around 1000 magnification.

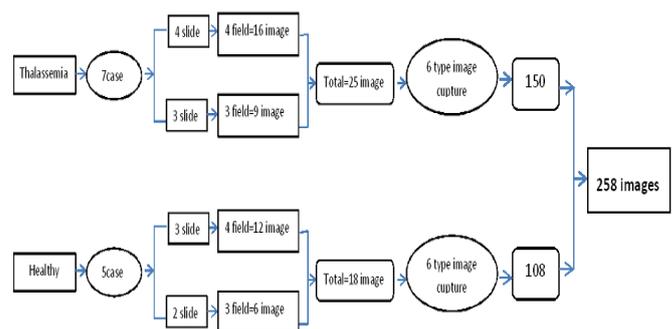
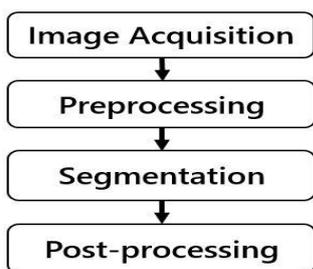


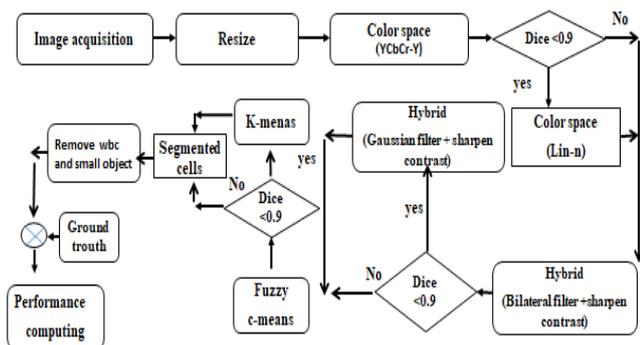
Fig.1: The image-capturing distribution

2.2. Image analysis Process

This method is separated into four essential steps: image acquisition, preprocessing, segmentation, and post-processing, as illustrated in Fig. 2A, and the proposed algorithm in Fig. 2B.



A



B

Fig.2.A) Describes the proposed algorithm's main steps, B) Flowchart for the proposed method.

2.2.1. Image Acquisition

Blood film preparation requires a slide, a tube, and a blood spreader. On a base slide, a blood droplet is inserted. When a spreader slide is moved backward at an angle of 30°–45° to the blood base slide, the blood is spread evenly across the slide (Dybas et al., 2022).

The prepared blood smear is dried using an air dryer, and the stained dried smearing is fixed with

methanol. Then, it is stained using Leishman stain, used for staining purposes (Fig. 3). These stained slides are then examined using a microscope (Dybas et al., 2022).

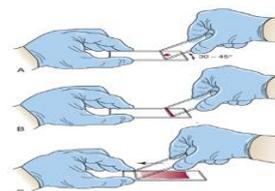


Fig. 3: Wedge technique of making a peripheral blood smear (Tyas et al., 2022).

The images were confirmed by an experienced hematologist working at Raparin Hospital-Kurdistan/Iraq for all microscopic blood smear images. Images were captured under nearly the same lab conditions (Biophysics Research lab /Education College / Salahaddin University – Erbil) for the yellow filter and the one without a filter, with different intensities (see Fig. 4).

2.2.2. Preprocessing

The image is resized from 3488 × 2616 pixels to 500 × 500 pixels to reduce the running time in the next stage. The preprocessing step was undertaken to refine and modify the image acquisition results such that the subsequent image processing stage produced more accurate segmentation results.

A color space is fundamentally an arrangement of hues (Gowda and Yuan, 2018).

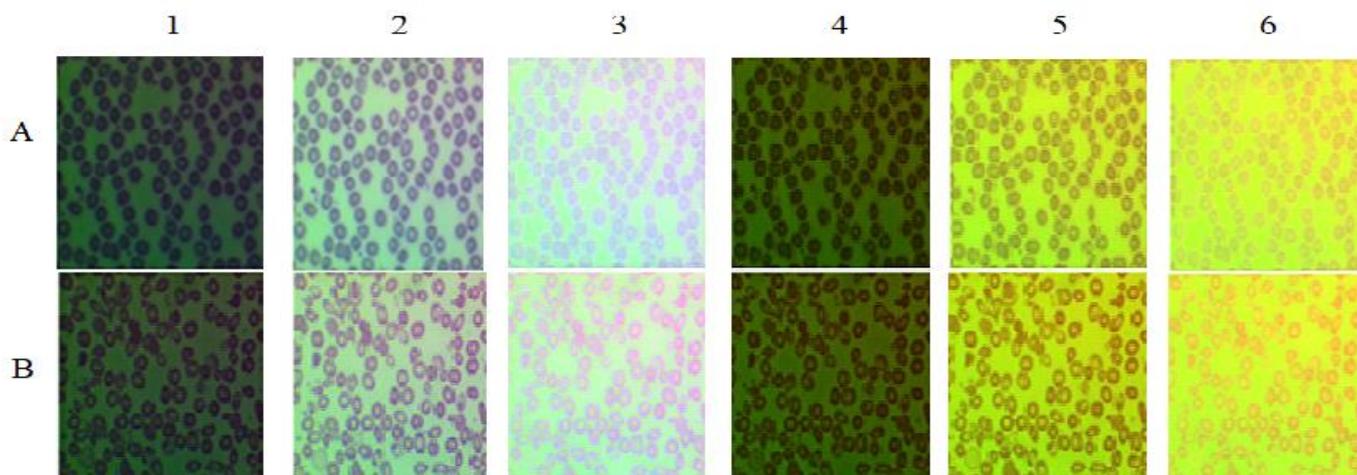


Fig. 4. A) The first row shows an image acquisition of healthy, and B) the second row shows an image acquisition of thalassemia patients. The 1 (without filter) and 4 (Yellow filter) columns for low intensity, 2 (without filter) and 5 (Yellow filter) columns for medium intensity, and 3 (without filter) and 6 (Yellow filter) columns for high intensity.

Converting color space from one type to another was one of the essential tasks in preprocessing. The application of color space to blood smears enhances and better segments their images segmented. Many factors negatively affect image quality, including improper camera use, poor lighting situation, etc. RGB is a helpful color model for computer graphics since the human visual system functions similarly to an RGB color space. It is represented by a mix of Red, Green, and Blue (Soloman, 2010). The input RGB image is simultaneously converted into the ten other color spaces; the HSV, XYZ, HSY, YCbCr, YIQ, YPbPr, CAT02 LMS, YDbDr, Lin, and YUV with their components, compared and the best of them (according to proposed performance) is utilized in the segmentation process. The HSV (Hue, Saturation, Value) is composed of the H-component (Hue, the color's position on a color wheel), The S-component (Saturation, the amount of hue), and the V-component (Value, the maximum value among the red, green, and blue components of a specific color) (Wahhab, 2015). The XYZ is have not a clear color, luminance of color, and amount of blue. The X-component is a mix of the image's three CIE RGB curves chosen to be nonnegative. Y-component is the luminance of color in the image (Wahhab, 2015).

The HSY is a color coordinate system categorizing colors in Hue, Saturation, and Luminosity (Abood et al., 2017). The YCbCr is Luminance, Chrominance Blue-component, Chrominance Red-component. As the Y-component increases, colors increase in brightness. The YIQ is Luminance; the amount of blue or orange, Amount of green or purple is known as NTSC (National Television Systems Committee) (Gowda and Yuan, 2018). The CIE CAT02 LMS is the linear transformation of XYZ using the MCAT02 chromatic adaptation matrix (Gandomkar et al., 2017). The YDbDr is composed of three components Y, Db, and Dr. Y is the luminance, and Db and Dr are the chrominance components (Sahdra and Kailey, 2012).

Image enhancement is improving image quality and enhancing the appearance of an image or a subset of an image to improve contrast or the visibility of specific characteristics and remove unwanted elements to enhance the accuracy of subsequent image analysis. Generally, various

techniques could be used for image enhancement: filtering and contrast (Oussama et al., 2017). The six filters used were: Gaussian, Wiener, Kalman, Non-local means, Median, and Bilateral. The three-contrast techniques were: Contrast-limited adaptive histogram equalization (CLAHE), Sharpen image and Adjust image intensity (AII).

2.2.3 Image Segmentation

Image Segmentation is the most crucial part of image processing. The purpose of segmentation is to facilitate image analysis by altering the image's representation. Segmentation divides an image into sections with comparable characteristics (Patgiri and Ganguly, 2021). In this study, the objective of segmenting a microscopic blood film is to isolate healthy and diseased RBCs from the background. Here, Fuzzy c-means and k-means are used as a hybrid segmentation technique.

Fuzzy C-means is a widely used soft clustering algorithm in which data points simultaneously belong to two or more clusters. In this approach, membership in each cluster center is allocated to each data point based on the distance between the cluster center and the data point (Patgiri and Ganguly, 2021). The k-means algorithm comprises two distinct phases. In the initial phase, the system calculates the k-centroid. In the second phase, each point is assigned to a cluster.

While there are several ways to define the distance from the nearest centroid, the Euclidean distance is one of the most widely used options. Once the grouping is complete, the new centroid of each is recalculated. A cluster's centroid is used to calculate a new Euclidean distance. It is estimated and assigned between each center and each data point; each cluster within a partition is defined by its constituent items and its centroid, the centroid of any cluster is the place to which the sum of the distances between all the objects in that cluster is minimal. Consequently, k-means is an iterative technique that minimizes the sum of distances between each item and the centroid of each cluster over all clusters (Abood et al., 2017).

2.2.4 Post processing

Even though the acquired image can distinguish the RBC region, noise, holes, small objects, and white blood cell count exist in the segmented image. This procedure is critical since

noise can significantly impair the system's ability to identify and classify the retrieved object appropriately. The morphological operation modifies objects' size, shape, structure, and connectivity in a binary image by changing their size, shape, structure, and connectivity. Using a structuring element and a set operator defined by Erosion and Dilation, Erosion is used to shrink and 'thin' objects in the image. In contrast, dilation is used to 'grow' and 'thicken' the image objects. The combination of the two operators can eliminate, break connections, clean the border, and fill in holes (Tomari et al., 2014). This study uses morphological operations to remove holes inside the WBC cell and small objects.

2.2.5 Segmentation Evaluation Parameters

Five Segmentation Evaluation Parameters are utilized in this study. There are requirements for calculating segmentation quality. All approaches require a ground truth image. The ground truth images were collected using a code and checked by a hematologist.

Dice similarity coefficient This metric is utilized by (Al-Hafiz et al., 2018) to evaluate segmentation accuracy based on the Dice coefficient. The Dice metric is an index of similarity determined using Eq.1:

$$Total\ Dice = \frac{2|X \cap Y|}{|X| + |Y|} \quad (1)$$

where X is represented by the segmented image and Y by the ground truth.

The Photoshop program obtained the ground truth image. The procedures are as follows: Make a black background with the magic wand tool and fill the dialogue box. Inverse select and make the foreground white using a fill dialogue box. Delete small or unwanted regions using the elliptical marquee tool.

Jaccard Similarity Coefficient (Jaccard) shows how different and similar the elements of a finite sample set are. As indicated in Eq.2, it is computed by dividing the number of elements common to crossing regions of X and Y by the sum of areas X and Y. (Prabha and Kumar, 2016).

$$Jaccard = \frac{|Y \cap X|}{|XUY|} \quad (2)$$

Probability Rand Index (PRI): During data clustering, the degree of similarity between two regions is determined. It compares the segmented

image to the ground truth to ensure consistent labeling. It counts a predetermined number of pixel pairs and averages the result across an image's ground truths. PRI is formally represented by Eq.3 (Zou et al., 2004).

$$PRI = \frac{A + B}{A + B + C + D} \quad (3)$$

A is the number of pairings of image elements belonging to the same subset in X and Y areas. B, the number of image element pairs in distinct subsets of area X and subsets of region Y. C, the number of image element pairs that belong to the same subset in area X but different subsets in region Y. D, the number of pairs of image elements that belong to distinct subsets in area X but the same subset in region Y.

Variation of Information (VOI): is a simple linear expression that measures the distance between two clusters of mutual information. In Eq.4, VOI computes the unpredictability of one segmentation based on its distance from another. See (Fig. 5) (Prabha and Kumar, 2016).

$$VOI(Y, X) = H(Y) + H(X) - 2I(X, Y) \quad (4)$$

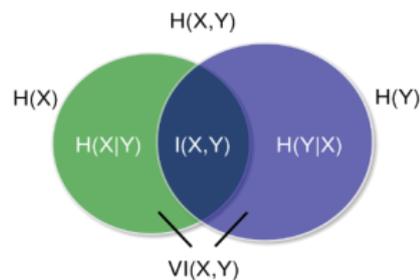


Fig.5: Variation of information parameters.

The Global Consistency Error (GCE) specifies the degree to which one segmentation differs in granularity. Consistent segmentation decisions are made since they may reflect the same image segmented at various scales. Segmentation refers to the arrangement of pixels inside an image. Pixel groups constitute segments. The pixel is in a refining zone if one segment is a perfect subset of the other segment and the error is zero. If there is no Subset relationship in the formula for GCE in Eq.5, the two zones will overlap unanticipatedly (Prabha and Kumar, 2016).

$$GCE = \frac{1}{n} \min \left\{ \sum_i E(q_1, q_2, p_i), \sum_i E(q_1, q_2, p_i) \right\} \quad (5)$$

Performance (i.e., number of correct segmenting with dice greater or equal to 0.9 per total number of images) in Eq. 6

$$\text{Performance} = \frac{\text{image of segmenting with dice} \geq 0.9}{\text{total number of images}} \times 100\% \quad (6)$$

RESULTS AND DISCUSSION

Image processing results included different steps (Figs.2B, 7, and 9). Each step was ranked based on performance and segmentation evolution parameters (Dice). Eleven other color spaces with their components were used to study their effects on segmentation performance (see Fig.6). The best color space was the YCbCr-Y color space, with a commission of **81%**. The second and third essential color spaces were the **RGB-G** and **lin-n components** (**lin** means Linearize gamma-corrected RGB values), with a performance of **78.3%** and **72.5 %**, respectively. The color spaces with zero performance were: YPbPr-pr, CAT02-s, YIQ-Q, YDbDr-Dr, XYZ-Z, YUV-V, YCbCr-Cr, HSY-S, HSV, HSV-H, and HSV-S components, which did not display in the chart in (Fig.6).

The best color space in the study is the **YCbCr-Y component** (Fig.6), which has also been identified as the most effective in other investigations. The microscopic images have

problems in the RGB color space and have the best results in the YCbCr color space (Kumar and Singh, 2012, Sharif et al., 2012). Also, another benefit of YCbCr is clearly displaying the minute details of image dark areas without stretching. When the proper stretching is applied, the results clearly show the differences (Sharif et al., 2012). The second important color space was the RGB-G component, which was counted as effective in segmentation efficiency, as mentioned in other works (Tyas et al., 2022, Tyas et al., 2020). The RGB-G component color space, which is the most widely used and most suitable for computer graphics, also uses the HSV color space for its ability for human perception. The importance of HSV was also confirmed by Huang Li et al. for edge detection of blood cells (Li et al., 2020). However, in this study, HSV color space needed to be revised in RBC segmentation.

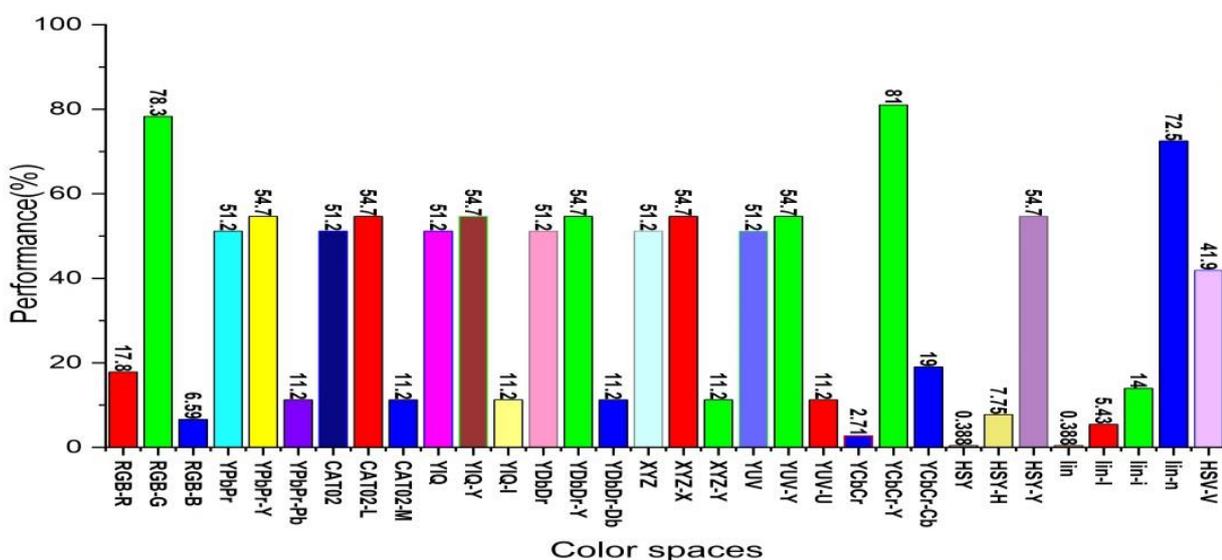


Fig.6: The effect of different color spaces with their components on the performance of proposed algorithm for image segmenting.

After choosing the color space, the **YCbCr-Y** component from (Fig. 7b) in the original image, this color space provides the image with the highest possible dice value, which is greater than (0.9). Once the best color space has been found, the enhancing steps include individual and hybrid to improve segmentation efficiency and performance.

The **Kalman filter** had a performance of **71.3%** among the five filters (Fig.8a), and the **sharp contrast** performance was **81.8%** among two contrast enhancers, which were more proper for segmentation. The implementation of the **Bilateral-filter-contrast-sharpening** enhancement was **82.9%**, which was better than the performance of the other hybrid techniques (Fig.

8b). The resulting image from the mentioned hybrid techniques is shown in Fig.7c.

A bilateral filter is an edge-preserving smoothing filter that filters coefficients determined based on both the spatial and range (pixel intensity) distances of the target image (Shirai et al., 2022), and that is important to have smoothing without degrading the boundary of RBC cells. Sharpening contrast is utilized as a hybrid enhancing method to improve visibility in low-contrast regions of RBC cells, and it may also be used for real-time picture processing. (Hoshi et al., 2022).

The most effective filters for adjusting image intensity (AII) were a **Kalman filter** with a performance of **79%** and a **CLAHE median filter** with a performance of **76.6%**, as depicted in (Fig. 8C and Fig. 8D).

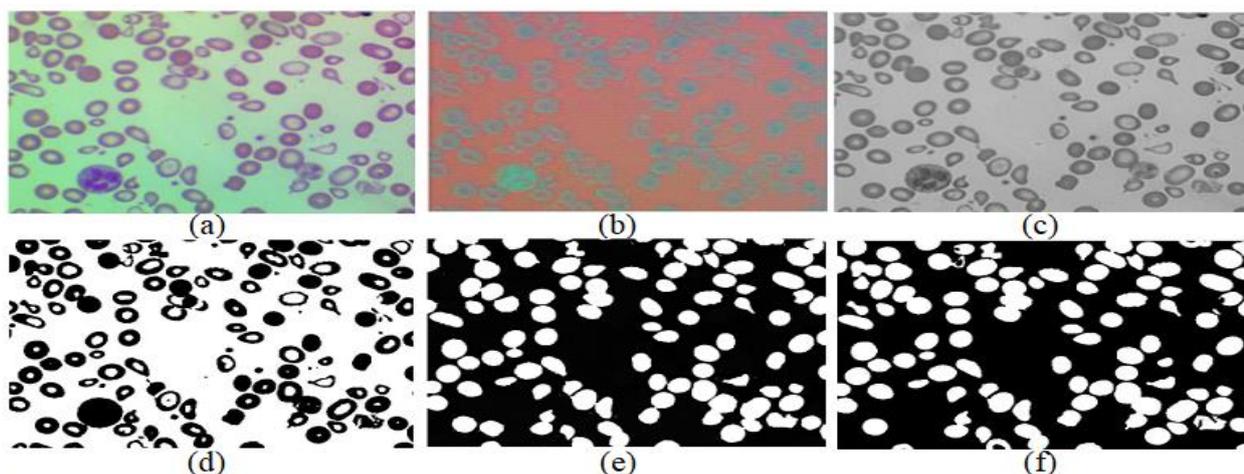


Fig. 7 Shows the original image of thalassemia patients (a), YCbCr-y color space (b), Hybrid (bilateral filter +sharpen contrast) (c), Fuzzy c-means of segmented (d), Remove WBC and small object (e), ground truth image (f).

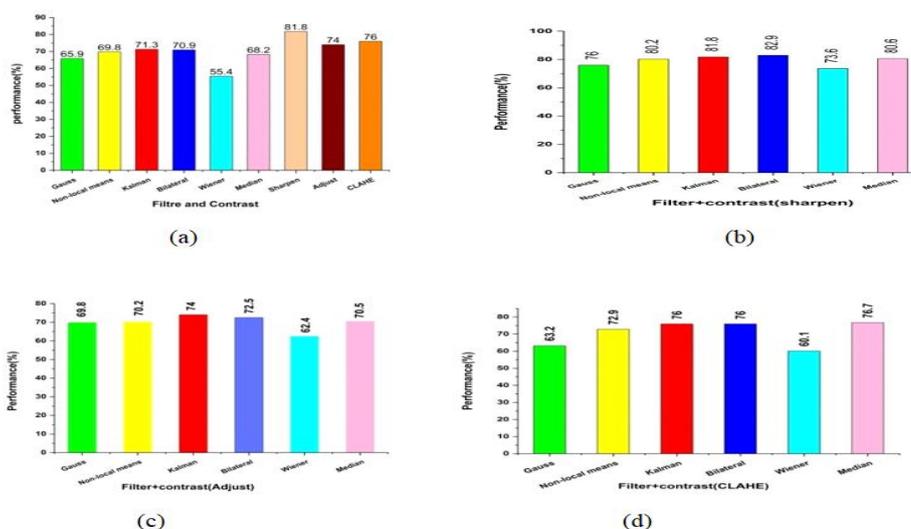


Fig. 8. Show an original image of thalassemia patients with filter and contrast; a) hybrid (filters+ contrast sharpening), b) hybrid (filters+ adjusting image intensity), c) hybrid (filter + contrast (Adjust)), d) hybrid (Filter+ contrast (CLAHE)).

After enhancing which qualified according to performance, the work comes to the segmentation stage. This step is essential for thalassemia detection due to specific erythrocyte morphology. Therefore, a suitable image segmentation strategy is required for separating morphological and textural in RBC cells (Tyas et al., 2020). This study employed fuzzy c-means, with a performance of **79.1%**. Fig.7d is the image obtained after the application of fuzzy c-means.

It targets distinguishing between RBC cells from the background. It can be used to separate the cytoplasm and central pallor, as done by (Kumar and Babulal, 2022).

To improve RBC cell segmentation performance, hybrid techniques were used to enhance performance of images in different

microscope light intensities with and without a yellow filter.

This work use lin (n-component) color space when dice score lower than 0.9 for each input image (Fig. 9a), and Fig. 9b shows the applied color space lin-n. The Gauss filter with sharpened contrast is also used for the same condition (Fig. 9c). Lastly, the k-means segmentation was used to improve segmentation quality in the same manner (Fig. 9d). The WBC and small objects were removed using the filter object function (i.e., bwareafilt function) and a sample output shown in Fig. 9e. In this work, the challenge of multi-cell selection was performed by Photoshop program as it was not realized by another researcher in any research (Fig. 9f).

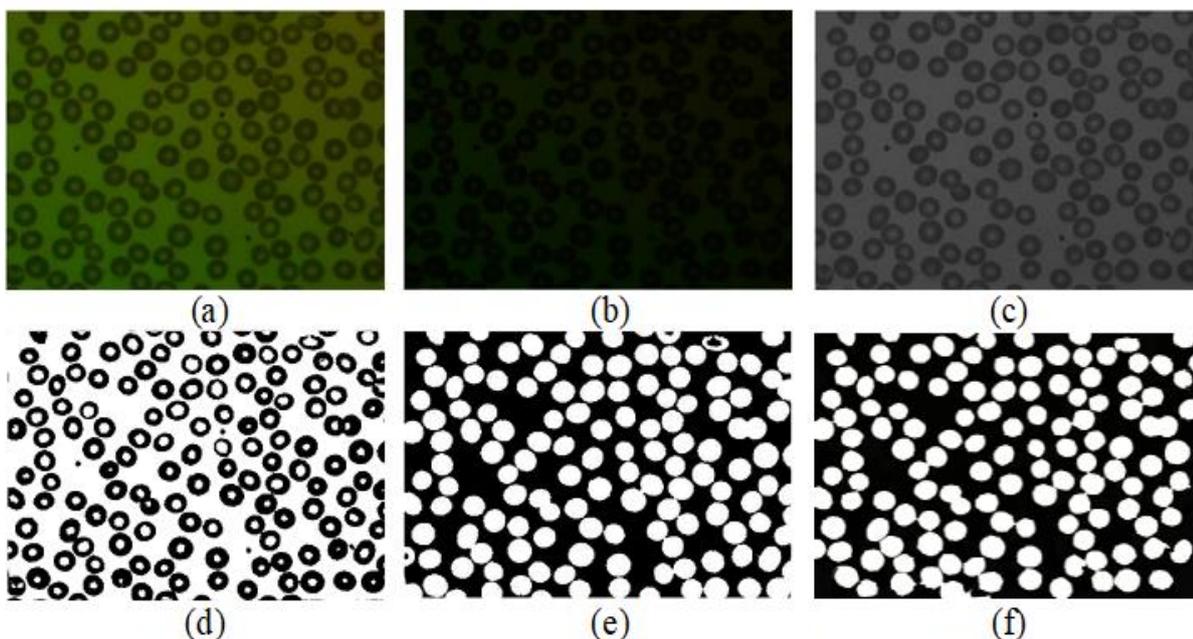


Fig. 9: Original image of healthy with a yellow filter (a) lin-n color space (b), Hybrid (Gaussian filter +sharpen contrast)(c), K -means of segmented (d), Remove WBC and small object (e), ground truth image (f)

The bar chart (Fig. 10) displays the individual and **hybrid of fuzzy c-means and k-means** performances. The highest performance belongs to the hybrid segmentation (**86.8%**) compared to the individual segmentation.

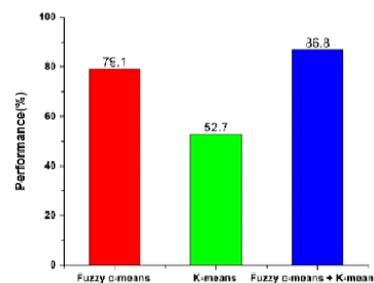


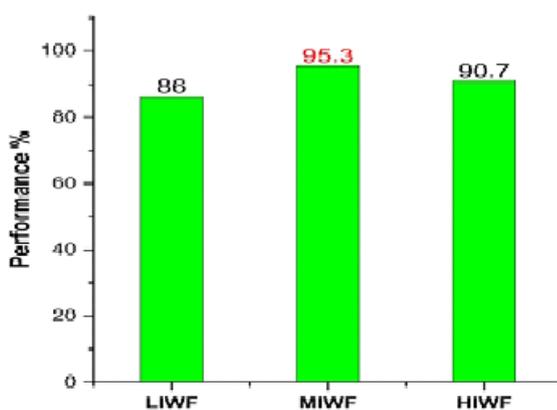
Fig. 10. Display different segmentation evaluation parameters by using: Fuzzy c-means, K-means, and Fuzzy c-means + K-means

One of the essential objectives of this study was to examine the effects of different microscope intensities with and without a yellow filter. Checking the influence of varying microscope light intensities and the use of a yellow filter on the captured image from the RBC blood film slide, as well as the image segmentation efficiency and performance, were essential objectives of this work.

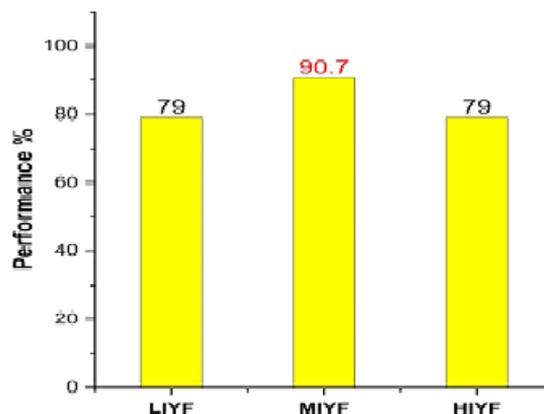
The medium intensity without filter (**MIWF**) gives **95.34%** performance (Fig.11a). The segmentation evaluation parameters (Table 1) for this condition were: DC (0.91), JSC (0.85), PRI (0.89), GCE (0.091), and VOI (0.55). The worst condition was for high intensity with a yellow filter (**HIYF**), that gives the lowest performance of 79% (Fig.11b). Their segmentation evaluation parameters were: DC (0.80), JSC (0.74), PRI (0.84), GCE (0.093) and VOI (0.643).

Table 1: the results of segmentation evaluation parameters using different microscope light intensities with and without a yellow filter

Parameters	LIWF	MIWF	HIWF	LIYF	MIYF	HIYF
PRI	0.87± 0.1	0.89± 0.06	0.88± 0.08	0.85± 0.12	0.85± 0.15	0.84± 0.12
GCE	0.094±0.033	0.091± 0.028	0.102± 0.038	0.096± 0.04	0.097± 0.06	0.093± 0.04
VOI	0.6± 0.19	0.55± 0.16	0.6± 0.21	0.642± 0.22	0.59± 0.16	0.643± 0.21
JSC	0.8± 0.22	0.85± 0.13	0.83± 0.14	0.76± 0.28	0.81± 0.22	0.74± 0.29
DC	0.86± 0.24	0.91±0.14	0.9± 0.11	0.81± 0.3	0.86±0.24	0.80 ± 0.3



(a)



(b)

Fig.11. physical parameters; a) without filter and with different intensities, b) filter and with different intensities.

Thus, the best microscope condition in the present work was for medium intensity without using a yellow filter. This condition can be changed according to the subject sample and environment intensity confirmed by other research

(Al-Dulaimi et al., 2018). Finally, the present work compared with several research displayed in Table 2.

Table 2: Comparing the accuracy values of the suggested approach to those of other segmentation evaluation researchers

Author, Year.	Al-Hafiz,et al., 2018	Aliyu et al., 2019	Chadha et al., 2020	The proposed method, 2022
No. of images	5	30	10	43
Segmentation	threshold and canny detector	Otsu threshold	Threshold	Hybrid Fuzzy c-means k-means
characteristic	87.9% (accuracy)	93 % (performance)	91.667% (performance)	95.34% (performance)

The compression result of the proposed algorithm with other works is displayed in table 2. (Al-Hafiz et al. (2018) obtained a dice similarity coefficient of 87.9% for RBC segmentation by threshold and canny detector. The authors Aliyu et al. (2019) received a performance of 93 % by using the Otsu threshold. Another study Chadha et al. (2020) used threshold segmentation and cell counting with a performance of 91.667%. The proposed algorithm achieves a performance of **95.34%**. Thus, the proposed work is distinguished by high performance, high-quality segmentation, and automated algorithm.

4. CONCLUSION

A novel technique for the automatic segmentation of red blood cells (RBC) from photographs of microscopic blood smears is presented. This study provides a novel mix of image processing techniques and extensive preprocessing to achieve superior segmentation performance. In particular, the best color space was the YCbCr-Y color space, with a performance of 81%; the performance of the Bilateral filter-wise contrast sharpening hybrid enhancement was 82.9%, and with fuzzy c-means segmentation, with a performance of 79.1%. Improve RBC cell

References

ABOOD, Z. M., KARAM, G. S. & HLUOT, R. E. Classification of red blood cells disease using fuzzy logic theory. 2017 International Conference on Current Research in Computer Science and Information Technology (ICCRIT), 2017. IEEE, 31-36.

segmentation performance obtained by using Lin-n color space, Gauss filter + sharpen contrast, and k-means segmentation, which raised performance to 86.8% for images in different microscope light intensities with and without a yellow filter. The selection of segmentation algorithms and their constituents is based on their resemblance to the ground-truth image using five assessment metrics. This study found that the best intensity for image slide capture utilizing a microscope was medium without using a yellow filter according to performance achievement (95.34%).

Acknowledgments

Thank you to the pathologist Prof Dr. Salah Abubaker Ali at Hawler Medical University and the hematologist Dr. Hewa A. Mustafa, at Children Raparin Teaching Hospital, for accepting me during my work. Thanks to Erbil-thalassemia center. Also, thanks to the Biophysics Research lab /Education College / Salahaddin University – Erbil.

Conflict of Interest: The authors declared no conflicts of interest.

AHMAD, I., ABDULLAH, S. N. H. S. & SABUDIN, R. Z. A. R. 2018. Morphological features analysis for erythrocyte classification in IDA and thalassemia. *International Journal of Advanced Computer Science and Applications*, 9.

AL-DULAIMI, K. A. K., BANKS, J., CHANDRAN, V., TOMEIO-REYES, I. & NGUYEN THANH, K. 2018. Classification of white blood cell types from microscope images: Techniques and challenges. *Microscopy science: Last approaches on*

- educational programs and applied research (Microscopy Book Series, 8), 17-25.*
- AL-HAFIZ, F., AL-MEGREN, S. & KURDI, H. 2018. Red blood cell segmentation by thresholding and Canny detector. *Procedia Computer Science*, 141, 327-334.
- ALIYU, H. A., RAZAK, M. A. A. & SUDIRMAN, R. Segmentation and detection of sickle cell red blood image. AIP conference proceedings, 2019. AIP Publishing LLC, 020004.
- CHADHA, G. K., SRIVASTAVA, A., SINGH, A., GUPTA, R. & SINGLA, D. 2020. An automated method for counting red blood cells using image processing. *Procedia Computer Science*, 167, 769-778.
- ÇIL, B., AYYILDIZ, H. & TUNCER, T. 2020. Discrimination of β -thalassemia and iron deficiency anemia through extreme learning machine and regularized extreme learning machine based decision support system. *Medical Hypotheses*, 138, 109611.
- DYBAS, J., ALCICEK, F. C., WAJDA, A., KACZMARSKA, M., ZIMNA, A., BULAT, K., BLAT, A., STEPANENKO, T., MOHAISSEN, T. & SZCZESNY-MALYSIAK, E. 2022. Trends in biomedical analysis of red blood cells—Raman spectroscopy against other spectroscopic, microscopic and classical techniques. *TrAC Trends in Analytical Chemistry*, 146, 116481.
- GANDOMKAR, Z., BRENNAN, P. C. & MELLOTHOMS, C. 2017. Determining image processing features describing the appearance of challenging mitotic figures and miscounted nonmitotic objects. *Journal of pathology informatics*, 8, 34.
- GOWDA, S. N. & YUAN, C. ColorNet: Investigating the importance of color spaces for image classification. Asian Conference on Computer Vision, 2018. Springer, 581-596.
- HOSHI, S., TASAKI, K., MARUO, K., UENO, Y., MORI, H., MORIKAWA, S., MORIYA, Y., TAKAHASHI, S., HIRAOKA, T. & OSHIKA, T. 2022. Improvement in Dacryoendoscopic Visibility after Image Processing Using Comb-Removal and Image-Sharpening Algorithms. *Journal of clinical medicine*, 11, 2073.
- KUMAR, P. & BABULAL, K. S. 2022. Hematological image analysis for segmentation and characterization of erythrocytes using FC-TriSDR. *Multimedia Tools and Applications*, 1-26.
- KUMAR, S. & SINGH, D. 2012. Colorization of Gray Scale Images in YCbCr Color Space Using Texture Extraction and Luminance Mapping. *IOSR Journal of Computer Engineering (IOSRJCE)*, 4, 27-32.
- LAVANYA, T. & SUSHRITHA, S. 2017. Detection of sickle cell anemia and thalassemia using image processing techniques. *International Journal of Innovative Science and Research Technology*, 2, 399-405.
- LI, H., ZHAO, X., SU, A., ZHANG, H., LIU, J. & GU, G. 2020. Color space transformation and multi-class weighted loss for adhesive white blood cell segmentation. *IEEE Access*, 8, 24808-24818.
- LIN, Y.-H., LIAO, K. Y.-K. & SUNG, K.-B. 2020. Automatic detection and characterization of quantitative phase images of thalassemic red blood cells using a mask region-based convolutional neural network. *Journal of Biomedical Optics*, 25, 116502.
- MARZUKI, N. I. B. C., BIN MAHMOOD, N. H. & BIN ABDUL RAZAK, M. A. 2017. Identification of thalassemia disorder using active contour. *Indonesian Journal of Electrical Engineering and Computer Science*, 6, 160-165.
- MOHAMED, M. & FAR, B. An enhanced threshold based technique for white blood cells nuclei automatic segmentation. 2012 IEEE 14th International Conference on e-Health Networking, Applications and Services (Healthcom), 2012. IEEE, 202-207.
- OUSSAMA, L., JAMIL, M., HAFIZAH, W. & AMBAR, R. 2017. Red blood cell image enhancement techniques for cells with overlapping condition. *Journal of Fundamental and Applied Sciences*, 9, 614-628.
- PATGIRI, C. & GANGULY, A. 2021. Adaptive thresholding technique based classification of red blood cell and sickle cell using naïve bayes classifier and k-nearest neighbor classifier. *Biomedical Signal Processing and Control*, 68, 102745.
- PELLEGRINO, R. V., TARROBAGO, A. C. & ZULUETA, D. L. B. Automated RBC Morphology Counting and Grading Using Image Processing and Support Vector Machine. 2021 IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), 2021. IEEE, 1-5.
- PRABHA, D. S. & KUMAR, J. S. 2016. Performance evaluation of image segmentation using objective methods. *Indian J. Sci. Technol*, 9, 1-8.
- PURWAR, S., TRIPATHI, R., RANJAN, R. & SAXENA, R. Classification of thalassemia patients using a fusion of deep image and clinical features. 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2021. IEEE, 410-415.
- RASHID, N. Z. N., MASHOR, M. Y. & HASSAN, R. Unsupervised color image segmentation of red blood cell for thalassemia disease. 2015 2nd International Conference on Biomedical Engineering (ICoBE), 2015. IEEE, 1-6.
- SADIQ, S., KHALID, M. U., ULLAH, S., ASLAM, W., MEHMOOD, A., CHOI, G. S. & ON, B.-W. 2021. Classification of β -Thalassemia Carriers From Red Blood Cell Indices Using Ensemble Classifier. *IEEE access*, 9, 45528-45538.
- SAHDRA, G. S. & KAILEY, K. S. 2012. Detection of Contaminants in Cotton by using YDbDr color space. *Int. J. Computer Technology & Applications*, 3, 1118-1124.
- SETSIRICHOK, D., PIROONRATANA, T., WONGSEREE, W., USAVANARONG, T., PAULKHAOLARN, N., KANJANAKORN, C., SIRIKONG, M., LIMWONGSE, C. & CHAIYARATANA, N. 2012. Classification of complete blood count and haemoglobin typing data by a C4.5 decision tree, a naïve Bayes classifier

- and a multilayer perceptron for thalassaemia screening. *Biomedical Signal Processing and Control*, 7, 202-212.
- SHARIF, J. M., MISWAN, M., NGADI, M., SALAM, M. S. H. & BIN ABDUL JAMIL, M. M. Red blood cell segmentation using masking and watershed algorithm: A preliminary study. 2012 international conference on biomedical engineering (ICoBE), 2012. IEEE, 258-262.
- SHIRAI, K., SUGIMOTO, K. & KAMATA, S.-I. 2022. Adjoint Bilateral Filter and Its Application to Optimization-based Image Processing. *APSIPA Transactions on Signal and Information Processing*, 11.
- SOLOMAN, S. 2010. *Sensors and control systems in manufacturing*, McGraw-Hill Education.
- TOMARI, R., ZAKARIA, W. N. W., JAMIL, M. M. A., NOR, F. M. & FUAD, N. F. N. 2014. Computer aided system for red blood cell classification in blood smear image. *Procedia Computer Science*, 42, 206-213.
- TYAS, D. A., HARTATI, S., HARJOKO, A. & RATNANINGSIH, T. 2020. Morphological, Texture, and Color Feature Analysis for Erythrocyte Classification in Thalassemia Cases. *IEEE Access*, 8, 69849-69860.
- TYAS, D. A., RATNANINGSIH, T., HARJOKO, A. & HARTATI, S. The classification of abnormal red blood cell on the minor thalassemia case using artificial neural network and convolutional neural network. *Proceedings of the International Conference on Video and Image Processing*, 2017. 228-233.
- TYAS, D. A., RATNANINGSIH, T., HARJOKO, A. & HARTATI, S. 2022. Erythrocyte (red blood cell) dataset in thalassemia case. *Data in Brief*, 41, 107886.
- WAHHAB, H. T. A. 2015. *Classification of acute leukemia using image processing and machine learning techniques*. University of Malaya.
- ZOU, K. H., WARFIELD, S. K., BHARATHA, A., TEMPANY, C. M., KAUS, M. R., HAKER, S. J., WELLS III, W. M., JOLESZ, F. A. & KIKINIS, R. 2004. Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports. *Academic radiology*, 11, 178-189.