

PAPER • OPEN ACCESS

Acute lymphoblastic leukemia image segmentation based on modified HSV model

To cite this article: Fallah H Najjar *et al* 2023 *J. Phys.: Conf. Ser.* **2432** 012020

View the [article online](#) for updates and enhancements.

You may also like

- [Compared growth mechanisms of Zn-polar ZnO nanowires on O-polar ZnO and on sapphire](#)
G Perillat-Merceroz, R Thierry, P-H Jouneau *et al.*
- [Impact of a van der Waals interface on intrinsic and extrinsic defects in an MoSe₂ monolayer](#)
Carlos J Alvarez, Minh Tuan Dau, Alain Marty *et al.*
- [Childhood leukaemia risks: from unexplained findings near nuclear installations to recommendations for future research](#)
D Laurier, B Grosche, A Auvinen *et al.*



245th ECS Meeting
San Francisco, CA
May 26–30, 2024




PRiME 2024
Honolulu, Hawaii
October 6–11, 2024

Bringing together industry, researchers, and government across 50 symposia in electrochemistry and solid state science and technology

Learn more about ECS Meetings at
<http://www.electrochem.org/upcoming-meetings>

 Save the Dates for future ECS Meetings!

Acute lymphoblastic leukemia image segmentation based on modified HSV model

Fallah H Najjar^{1*}, Kifah T Khudhair², Zaid Nidhal Khudhair^{3,4}, Haneen H Alwan² and Ameer Al-khaykan⁵

¹Department of Computer System Techniques, Technical Institute of Najaf, Al-Furat Al-Awsat Technical University, Iraq

²Department of Information Technology, Technical College of Management, Al-Furat Al-Awsat Technical University, Iraq

³Computer Techniques Engineering Department, Faculty of Information Technology, Imam Jaafar Al-Sadiq University, Baghdad, Iraq

⁴Faculty of Engineering, School of Computing, University Technology of Malaysia, Johar Bahru, Malaysia

⁵Air conditioning and Refrigeration Techniques Engineering Department, Al-Mustaqbal University College, 51001, Hillah, Babylon, Iraq

*Corresponding Author E-mail: fallahnajjar@atu.edu.iq

Abstract. Image segmentation is a critical step in computer-aided diagnosis that could speed up Leukemia detection. Leukemia is a cancer of the blood that has a reputation for being particularly lethal. Based on the immunohistochemical method, the leukocytes can be manually counted in a stained peripheral blood smear image to detect Acute Lymphoblastic Leukemia (ALL). Regrettably, the manual diagnosis process takes about 3 to 24 hours to complete, which is insufficient. This paper introduced a new and straightforward ALL image segmentation approach based on color image transformation. First, Leukemia, ALL-IDB1, ALL-IDB2, and ALL image datasets were used in this paper. The Leukemia dataset includes 208 ALL-IDB1 and ALL-IDB2 images, while The ALL dataset has 3256 images. Next, we use the HSV model to transform ALL images. In addition, we modified the HSV model by pre-processing the saturation channel for better results. Then, the pre-processed images were segmented based on a fixed threshold. After that, various metrics are utilized to measure the output of the proposed method. Finally, the proposed methodology is compared to currently used benchmarks. The proposed method outperforms previous approaches regarding accuracy, specificity, sensitivity, and time. In addition, results show that the proposed technique improves performance measures significantly.

1. Introduction

Determining cancer's stage (extent) is necessary for most cancer types. The size of the tumor and the scope to which cancer has spread control the location. However, this can be useful for determining a patient's prognosis and treatment. In contrast, Acute Lymphocytic Leukemia (ALL) does not typically form tumors. Instead, it usually affects all of the bone marrow in the body, which can spread to other organs, such as the liver and spleen. So, unlike most other cancers, ALL is not staged [1]. In 2022, the American Cancer Society reported, based on gender, that the estimated number of new leukemia is 80650 cases (35810 male and 24840 female), while the estimated number of death is 24000 cases (14020



male and 9980 female) [1]. In the medical community, leukemia is always seen as a life-threatening disease [2-5].

However, White Blood Cells (WBC) or leukocytes play an essential role in diagnosing various diseases, including leukemia; as a result, hematologists can benefit from obtaining information regarding these cells [6]. For example, the ALL analysis can be carried out through manually counting or machine-based procedures. ALL is typically detected by measuring the WBC in an immunohistochemical (IHC) image of a stained peripheral blood smear in a medical setting. Regrettably, the manual diagnosis process takes about 3 to 24 hours to complete [7]. Thus, an incorrect leukocyte cell count can result in an under or over-diagnosis of ALL [5]. Because of this, early detection of ALL using image processing and machine learning techniques, such as segmentation, can reduce the toxicity of cancer patients [6]. Medical imaging must be supplemented with a human physiology and anatomy database to diagnose abnormalities accurately. Images used in medical diagnostics visually represent the human body's internal structures, allowing pathologists to assess the severity of disease and tumor growth [8-10].

Several studies have used different imaging techniques to detect cancer, such as X-rays, tomography, MRI, and microscopy images [8, 11]. This study considered only microscopy images of blood cells. Image processing and machine learning techniques are practical in detecting acute lymphoblastic leukemia cells. However, segmenting leukocyte cells is the most critical step in determining these instruments' nucleus characteristics [12, 13]. Therefore, image segmentation is a vital component of any system used to diagnose ALL automatically. Several color space techniques, such as RGB, YCbCr, CMYK, and LAB, are currently used to partition microscopic images.

Consequently, this study introduces a new ALL-segmentation technique based on a modified HSV model. The main contribution of this research is an accurate and efficient segmentation of ALL images using a robust algorithm. On the other hand, one of the most critical pre-processing steps was converting color space from one type to another. Hence, images can be segmented using a variety of different techniques.

2. Related work

Several image segmentation methods have been proposed, including deep learning [14], Fuzzy C-means [15], K-means clustering, Otsu's thresholding [16], Watershed [17], deep learning, and other algorithms [18]. For example, in [19], the authors proposed a new segmentation approach using the k-means cluster and morphological operation on 642 images and got an overall segmentation accuracy of 97.47. Also, [20] used the K-mean clustering, but this time with the Zack algorithm; unfortunately, the authors used only seven images to evaluate the proposed method and got an overall segmentation accuracy of 96.60. Similarly, [21-25] proposed an ALL segmentation approach based on hybrid histogram-based soft covering rough k-means clustering. In addition, they used several features extracted from 350 images derived from the Gray Level Cooccurrence matrix and got an overall segmentation accuracy above 90%. However, the following sections go into more detail about the research. Section 2 describes an overview of the proposed model. Then, section 3 discusses the findings and conclusions of the analysis. Finally, section 4 concludes with a summary of the article's main points.

3. Methodology

In this study, our proposed method is divided into image pre-processing and image segmentation. In pre-processing, we convert ALL-IDB1, ALL-IDB2, and ALL images to the HSV (Hue, Saturation, Value) color space. As for the image segmentation part, we proposed a fixed threshold to segment the converted images. Finally, we will look at how well each dataset was segmented.

3.1. Image pre-processing

This subsection will demonstrate the HSV model as it relates to our pre-processing. Nevertheless, first, we chose to work in the HSV (Hue, Saturation, and Value) color space because it enables us to conduct our analysis of blood cell images in a manner that is analogous to human visual perception system works. Hue (H) refers to the overarching color that humans can perceive. Saturation (S) is the measure of the

amount of white light varying in hue. Finally, the Value (V) is the brightness or intensity of the color. In this study, the saturation channel is first adjusted by the following equation to increase the amount of light.

$$\text{adjusted } S = (2 \times S) \quad 1$$

Then, H, adjusted S, and V channels are recombined. Now, the modified images are ready for the next step. **Figure 4**, **Figure 5**, and **Figure 6** illustrate the results of pre-processing step on the ALL- dataset, ALL-IDB1, and ALL-IDB2, respectively.

3.2. Segmentation

After applying our proposed pre-processing on the ALL, ALL-IDB1, and ALL-IDB2 datasets, we used our proposed segmentation to extract the Region of Interest (ROI) in images. By using the following equations.

$$\text{Segmented image} = \begin{cases} \text{if adjusted } S > 0.7 & \text{keep the ROI} \\ \text{otherwise} & 0 \end{cases} \quad 2$$

3.3. Segmentation performance

The performance of segmentation serves as a benchmark for determining whether the proposed method is competent. Comparing the results of the segmented image with the ROI to the ground truth image is how the quality of the segmented image is determined. This comparison is performed based on the similarity of the pixels. Formulas for accuracy (Acc), specificity (SPE), and sensitivity (SEN) can be found as follows:

$$\text{Acc} = ((TP + TN)/(TP + TN + FP + FN)) \times 100 \quad 3$$

$$\text{SPE} = (TN/(TN + FP)) \times 100 \quad 4$$

$$\text{SEN} = (TP/(TP + FN)) \times 100 \quad 5$$

3.4. Dataset

The lack of a consistent number of images used from the publicly available datasets was the most significant problem encountered while comparing our proposed method. Previous researchers used publicly available images to test their systems, preventing them from reaching their results with other proposed methods. Hence, we used two available datasets in this paper: the Acute Lymphoblastic Leukemia (ALL) image dataset [26], see Figure 1, and Leukemia Dataset [27], see Figure 2. The ALL dataset consists of four classes (Hematogone (504), Early Pre-B (985), Pre-B (963), and Pro-B ALL (804)) with in total of 3256 images.

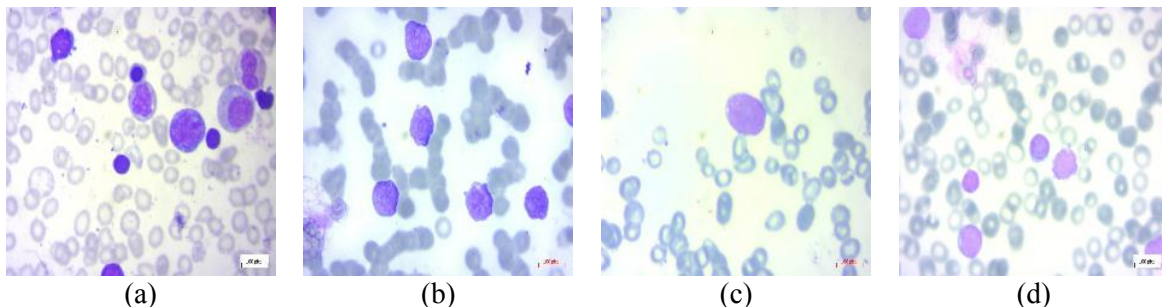


Figure 1. samples of ALL dataset images: (a) Hematogone, (b) Early Pre-B, (c) Pre-B, (d) Pro-B

In contrast, the Leukemia dataset consists of ALL-IDB1 (108) and ALL-IDB2 (260) with in total of 368 images.

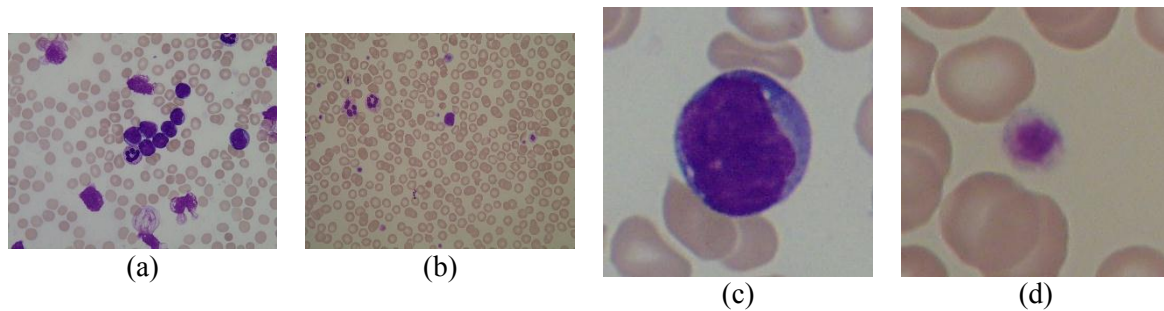


Figure 2. samples of Leukemia dataset images: (a, b) ALL-IDB1, (c, d) ALL-IDB2

Next, we will demonstrate the proposed method's algorithm and structure. **Table 1** shows an overview of the algorithm of the proposed method.

Table 1. Algorithm of the proposed method

Algorithm
Input: RGB image.
Output: Segmented image.
Step1: Read the image (RGB).
Step2: Convert RGB to HSV.
Step3: Extract the H, S, and V channels from the HSV image.
Step4: Apply equation (1) to the S channel to get Adjusted S.
Step5: Apply equation (2) to the Adjusted S.
Step6: Recombine the H, Adjusted S, and V to get the segmented image.
Step7: Calculate the segmentation performance based on the ground truth and the segmented image using equations (3), (4), and (5).

Furthermore, **Figure 3** shows an overview of the flowchart of the proposed method.

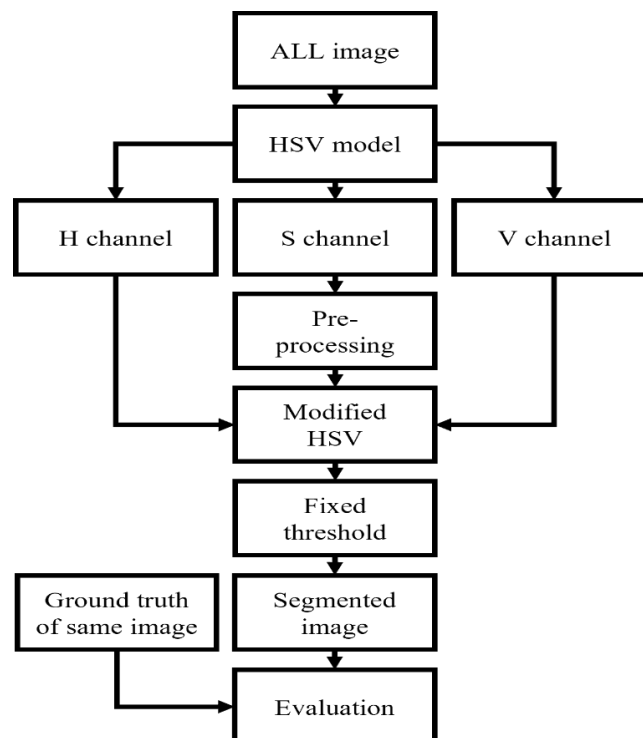


Figure 3. Flowchart of the proposed method

4. Results and discussion

This section will discuss and elaborate on the findings of this study. Image pre-processing and thresholding are both used in the leukemia segmentation process. This study tested the proposed segmentation method on 3624 images taken from the ALL, ALL-IDB1, and ALL-IDB2 datasets.

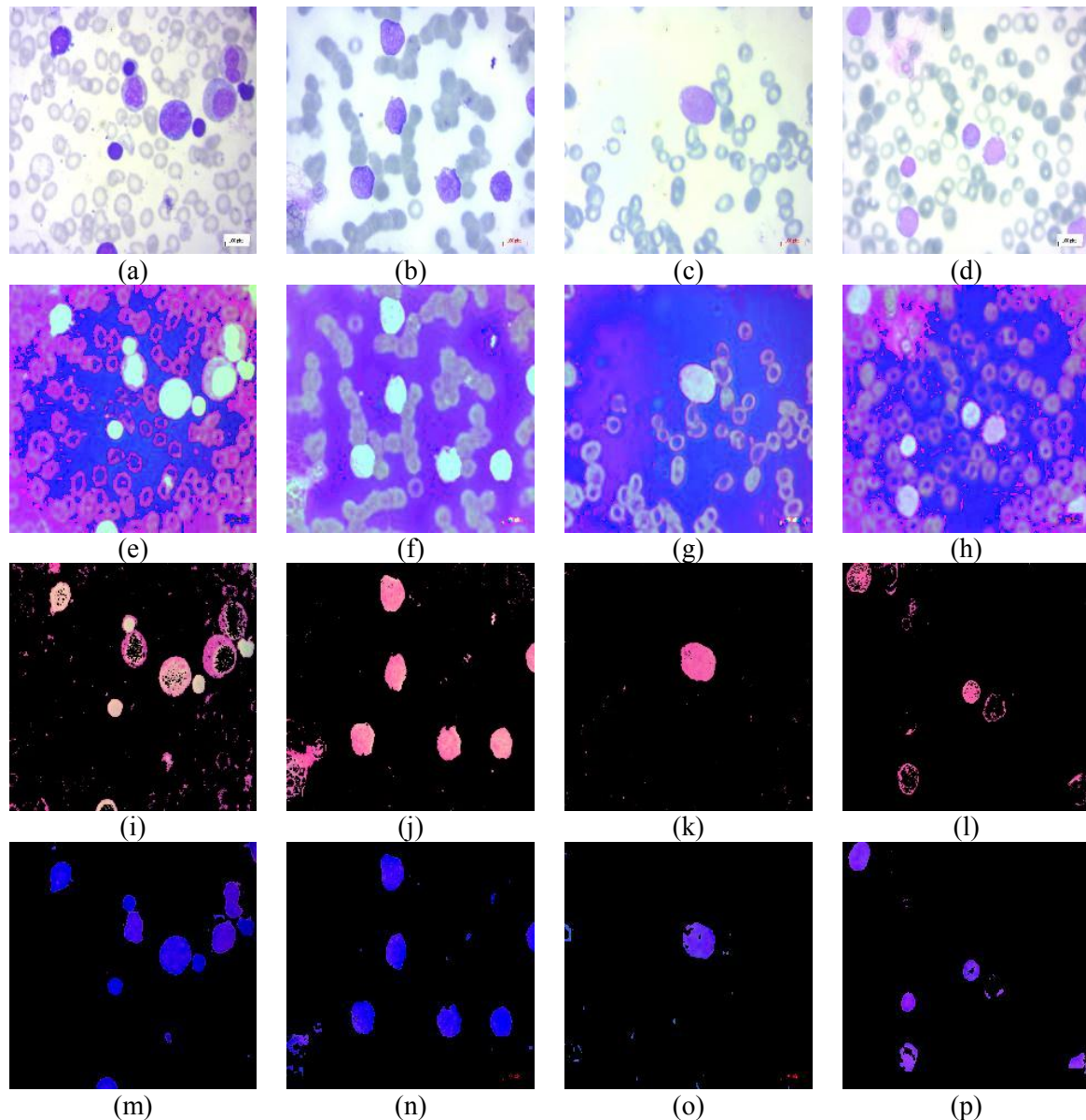


Figure 4. Proposed method results of the ALL dataset: (a)-(d) original images, (e)-(h) modified HSV, (i)-(l) ground truth, (m)-(p) segmented images.

For the ALL dataset, Figure 4 (a)-(d) shows the original image for each class. Next, Figure 4 (e)-(h) depicts the process of the modified HSV. Then, Figure 4 (i)-(l) displays the ground truth of the original images. Finally, in Figure 4 (m)-(p), the images have been segmented using the threshold we proposed. For the ALL-IDB1 dataset, Figure 5 (a) shows the original image for each class. Next, Figure 5 (b) displays the HSV images. Then, Figure 5 (c) depicts the process of the modified HSV. Finally, in Figure 5 (d), the images have been segmented using the threshold we proposed.

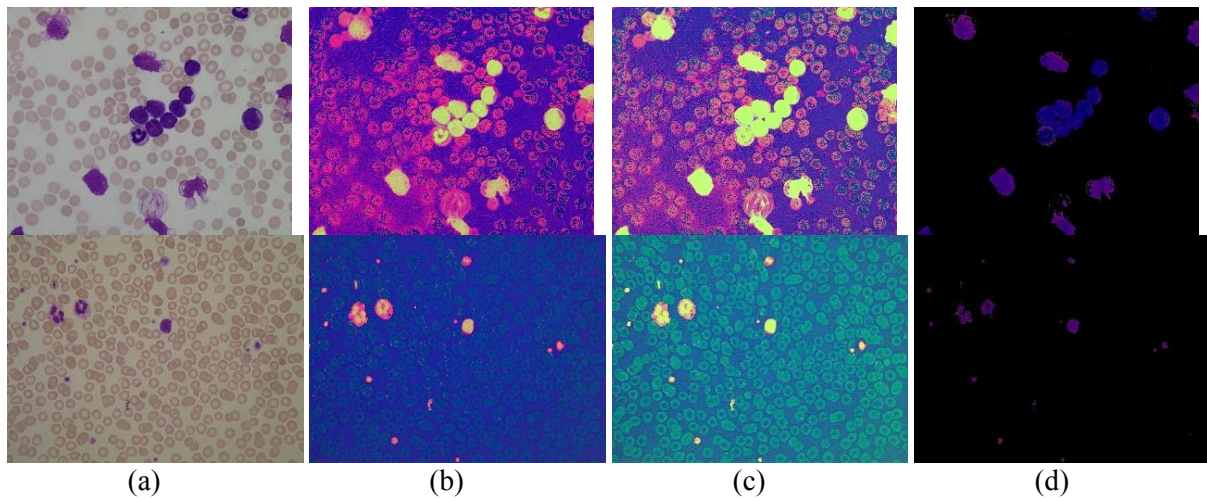


Figure 5. Proposed method results of the ALL-IDB: (a) original images, (b) HSV, (c) modified HSV, (d) segmented images.

For the ALL-IDB2 dataset, Figure 6 (a) shows the original image for each class. Next, Figure 6 (b) displays the HSV images. Then, Figure 6 (c) depicts the process of the modified HSV. Finally, in Figure 6 (d), the images have been segmented using the threshold we proposed.

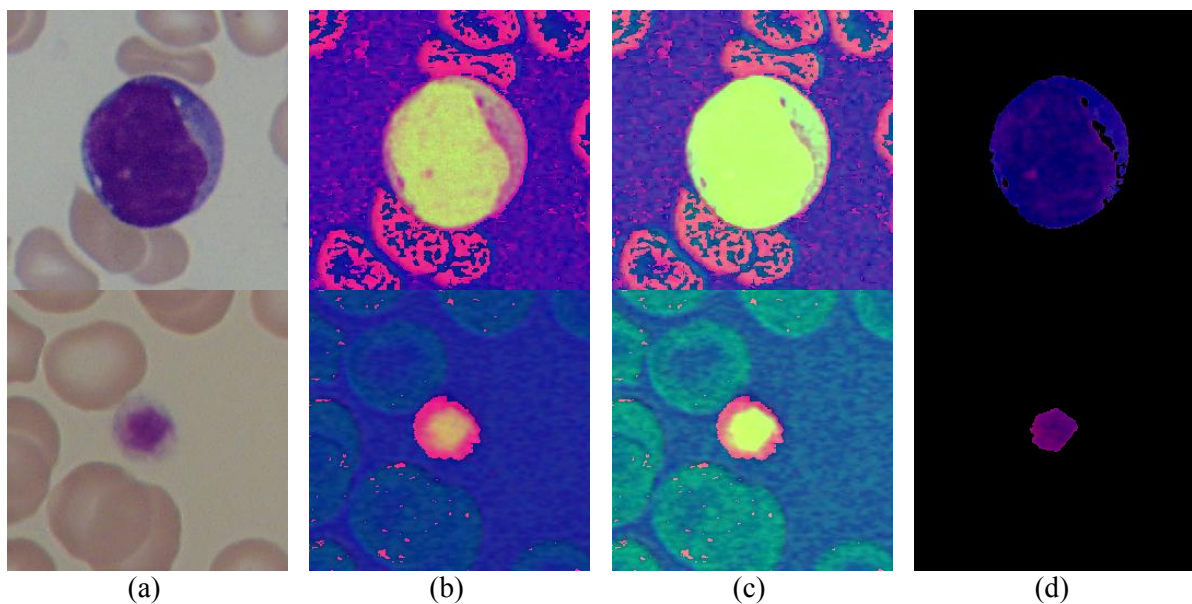


Figure 6. Proposed method results of the ALL-IDB2: (a) original images, (b) HSV, (c) modified HSV, (d) segmented images.

Furthermore, since the ground truth images for the ALL-dataset are provided, we compute three segmentation metrics for all dataset images Acc, SEN, and SPE. According to Acc, SEN, and SPE percentage, the results of segmentation performance are shown in Table 2.

However, to implement the proposed method, MATLAB (R2022a) and an Intel Core i7-11800H processor (2.30 GHz, 16 CPUs), 16 GB DDR4 RAM, and an NVIDIA GeForce RTX 3050 TI GPU were used in conjunction with Windows 11.

Table 2. Segmentation performance results for the ALL dataset images

Class	Acc (%)	SEN (%)	SPE (%)
Hematogone	95.82	86.50	97.83
Early Pre-B	98.17	82.12	99.42
Pre-B	92.02	81.95	91.76
Pro-B ALL	98.51	86.43	99.19

Additionally, we calculate the execution time of the proposed method for all images in both datasets. Figure 7 displays the execution time for each class in the ALL dataset.

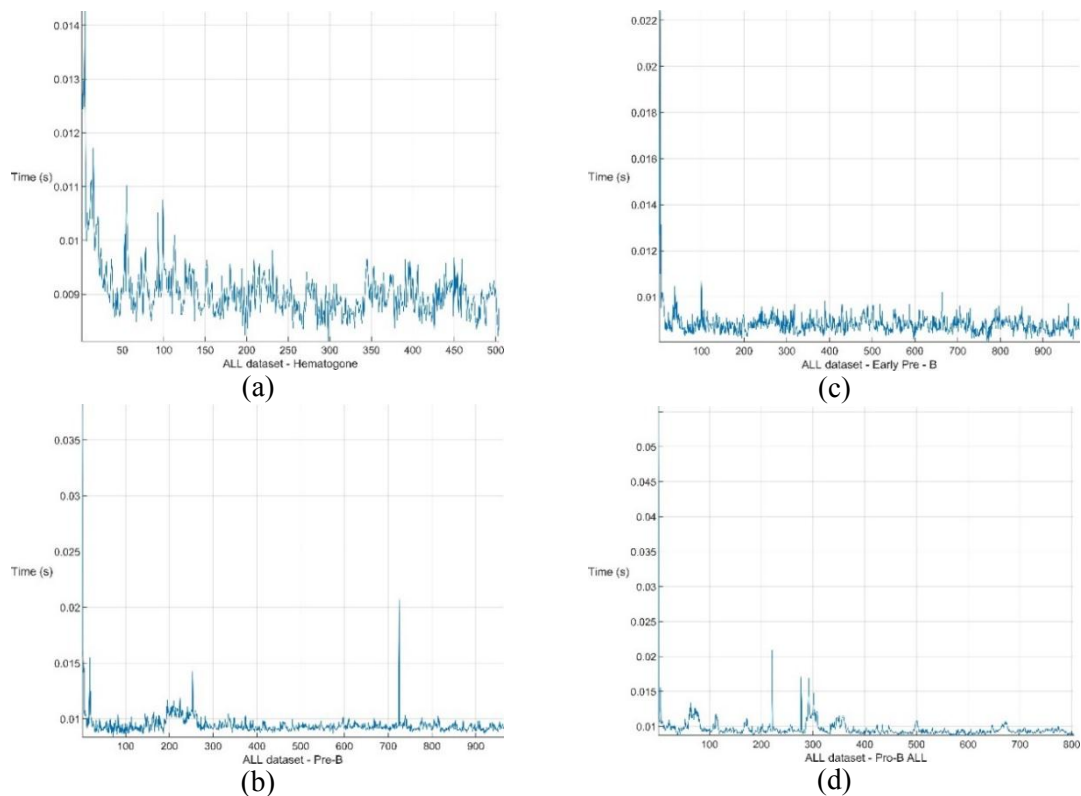


Figure 7. Processing time: (a) Hematogone (504 images), (b) Early Pre-B (985 images), (c) Pre-B (963 images), (d) Pro-B ALL (834 images).

Figure 8 displays the execution time for each ALL-IDB1 and ALL-IDB2 datasets.

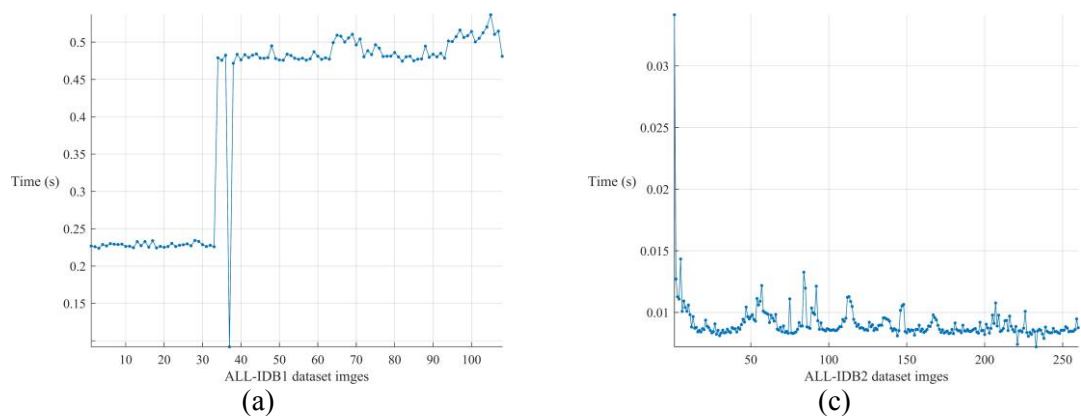


Figure 8. Processing time: (a) ALL-IDB1 (108 images), (b) ALL-IDB2 (260 images).

5. Conclusions

In this paper, we chose to work in the HSV color space because it enables us to conduct our analysis of blood cell images in a manner that is analogous to the way the human visual perception system works. First, however, we presented a new, straightforward method for segmenting ALL images using color image transformation. There are 208 ALL-IDB1 and ALL-IDB2 images in the Leukemia dataset and 3256 images in the ALL dataset. After that, we applied the HSV model to all images. Finally, to improve our results, we pre-processed the saturation channel in the HSV model. This was followed by a fixed-threshold segmentation of the previously pre-processed images. In the next step, various metrics are used to evaluate the proposed method's performance. Finally, a comparison is made between the proposed approach and the standard currently in use. The proposed method outperforms previous methods in terms of accuracy, specificity, sensitivity, and time. Furthermore, the results show that the proposed technique significantly enhances performance metrics.

References

- [1] "American Cancer Society." <https://cancerstatisticscenter.cancer.org/#!/cancer-site/Leukemia> (accessed 9 - 5 - 2022).
- [2] N. M. Deshpande, S. S. Gite, and R. Aluvalu, "A brief bibliometric survey of leukemia detection by machine learning and deep learning approaches," *Lib. Philo. Pract.*, vol. 4569, 2020.
- [3] Salim, A. A., Bidin, N., Lafi, A. S., & Huyop, F. Z. (2017). Antibacterial activity of PLAL synthesized nanocinnamon. *Materials & Design*, 132, 486-495.
- [4] Salim, A. A., Bidin, N., & Ghoshal, S. K. (2018). Growth and characterization of spherical cinnamon nanoparticles: Evaluation of antibacterial efficacy. *LWT*, 90, 346-353.
- [5] Salim, A. A., Ghoshal, S. K., Krishnan, G., & Bakhtiar, H. (2020). Tailored fluorescence traits of pulse laser ablated Gold-Cinnamon nanocomposites. *Materials Letters*, 264, 127335.
- [6] L. H. Vogado, R. M. Veras, F. H. Araujo, R. R. Silva, and K. R. Aires, "Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification," *Engineering Applications of Artificial Intelligence*, vol. 72, pp. 415-422, 2018.
- [7] I. Naz, N. Muhammad, M. Yasmin, M. Sharif, J. H. Shah, and S. L. Fernandes, "Robust discrimination of leukocytes protuberant types for early diagnosis of leukemia," *Journal of Mechanics in Medicine and Biology*, vol. 19, no. 06, p. 1950055, 2019.
- [8] J. Rawat, H. Bhaduria, A. Singh, and J. Virmani, "Review of leukocyte classification techniques for microscopic blood images," in *2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom)*, 2015: IEEE, pp. 1948-1954.
- [9] G. Jothi, H. H. Inbarani, A. T. Azar, and K. R. Devi, "Rough set theory with Jaya optimization for acute lymphoblastic leukemia classification," *Neural Computing and Applications*, vol. 31, no. 9, pp. 5175-5194, 2019.
- [10] M. A. Mohammed, M. K. Abd Ghani, R. I. Hamed, and D. A. Ibrahim, "Review on Nasopharyngeal Carcinoma: Concepts, methods of analysis, segmentation, classification, prediction and impact: A review of the research literature," *Journal of Computational Science*, vol. 21, pp. 283-298, 2017.
- [11] S. R. Waheed, M. H. Alkawaz, A. Rehman, A. S. Almazyad, and T. Saba, "Multifocus watermarking approach based on discrete cosine transform," *Microscopy Research and Technique*, vol. 79, no. 5, pp. 431-437, 2016.
- [12] Salim, A. A., Bakhtiar, H., Ghoshal, S. K., & Huyop, F. (2020). Customised structural, optical and antibacterial characteristics of cinnamon nanoclusters produced inside organic solvent using 532 nm Q-switched Nd: YAG-pulse laser ablation. *Optics & Laser Technology*, 130, 106331.
- [13] Salim, A. A., Ghoshal, S. K., & Bakhtiar, H. (2021). Tailored morphology, absorption and bactericidal traits of cinnamon nanocrystallites made via PLAL method: Role of altering laser fluence and solvent. *Optik*, 226, 165879.

- [14] S. R. Waheed, M. M. Adnan, N. M. Suaib, and M. S. M. Rahim, "Fuzzy logic controller for classroom air conditioner," in *Journal of Physics: Conference Series*, 2020, vol. 1484, no. 1: IOP Publishing, p. 012018.
- [15] S. S. Al-jaboriy, N. N. A. Sjarif, S. Chuprat, and W. M. Abdulllah, "Acute lymphoblastic leukemia segmentation using local pixel information," *Pattern Recognition Letters*, vol. 125, pp. 85-90, 2019.
- [16] M. M. Adnan, M. S. M. Rahim, K. Al-Jawaheri, M. H. Ali, S. R. Waheed, and A. H. Radie, "A survey and analysis on image annotation," in *2020 3rd International Conference on Engineering Technology and its Applications (IICETA)*, 2020: IEEE, pp. 203-208.
- [17] S. R. Waheed, N. M. Suaib, M. S. M. Rahim, M. M. Adnan, and A.A. Salim, "Deep Learning Algorithms-based Object Detection and Localization Revisited," in *Journal of Physics: Conference Series*, 2021, vol. 1892, no. 1: IOP Publishing, p. 012001.
- [18] H. Iilmadina and A. Arymurthy, "A New Image Segmentation of Leptomeningeal Metastasis in Leukemia Patients," in *Journal of Physics: Conference Series*, 2020, vol. 1577, no. 1: IOP Publishing, p. 012014.
- [19] E. Suryani, E. Asmari, and B. Harjito, "Image Segmentation of Acute Myeloid Leukemia Using Multi Otsu Thresholding," in *Journal of Physics: Conference Series*, 2021, vol. 1803, no. 1: IOP Publishing, p. 012016.
- [20] N. Daud, R. Raof, M. Osman, and N. Harun, "Segmentation Technique for Nucleus Detection in Blood Images for Chronic Leukaemia," in *Journal of Physics: Conference Series*, 2021, vol. 1755, no. 1: IOP Publishing, p. 012053.
- [21] A. Hashim and H. Mahdi, "Object Detection and Recognition Using Local Quadrant Pattern," *Journal of Kufa for Mathematics and Computer*, vol. 6, no. 2, 2019.
- [22] J. Su, S. Liu, and J. Song, "A segmentation method based on HMRF for the aided diagnosis of acute myeloid leukemia," *Computer methods and programs in biomedicine*, vol. 152, pp. 115-123, 2017.
- [23] N. Patel and A. Mishra, "Automated leukaemia detection using microscopic images," *Procedia Computer Science*, vol. 58, pp. 635-642, 2015.
- [24] M. G. Mehrad Aria, Davood Bashash, Hassan Abolghasemi, Farkhondeh Asadi, Azamossadat Hosseini. Acute Lymphoblastic Leukemia (ALL) image dataset [Online] Available: <https://www.kaggle.com/datasets/mehradaria/leukemia>
- [25] Salim, A. A., Ghoshal, S. K., Bakhtiar, H., Krishnan, G., & Sapingi, H. H. J. (2020, April). Pulse laser ablated growth of Au-Ag nanocolloids: Basic insight on physiochemical attributes. In *Journal of Physics: Conference Series* (Vol. 1484, No. 1, p. 012011). IOP Publishing.
- [26] Salim, A. A., Bidin, N., Ghoshal, S. K., Islam, S., & Bakhtiar, H. (2018). Synthesis of truncated tetrahedral cinnamon nanoparticles in citric acid media via PLAL technique. *Materials Letters*, 217, 267-270.
- [27] N. sharma. Leukemia Dataset [Online] Available: <https://www.kaggle.com/datasets/nikhilsharma00/leukemia-dataset>