DOI:10.9756/INTJECSE/V14I5.475 ISSN: 1308-5581 Vol 14, Issue 05 2022

An Advanced Low-cost Blood Cancer Detection System

Christo Ananth,
Department of Natural and Exact Sciences, Samarkand State University, Uzbekistan.
K. Jasmine Mystica,
Department of ECE, St. Joseph's College of Engineering, India.
M. Sridharan,
Department of Mathematics, PSNA College of Engineering and Technology, India.
S. Aaron James,
University of Technology and Applied Sciences, Ibri, Oman.
T. Ananth Kumar,
Assistant Professor, Computer Science and Engineering, IFET College of Engineering, India.

Abstract--- Blood leukemia can be diagnosed with greater precision and specificity when red blood cells are separated from young white blood cells. The only way to tell if someone has a blood disorder is to take photographs of their skin, calculate, shade, and measure them. According to the World Health Organization, leukemia is the fifth leading cause of death in the world. Early detection and identification of infection are critical to the treatment's success and outcome. It is hoped that the researchers will be able to identify and test cells that are linked to leukemia in this study. This disease can be classified according to whether or not it has immature cells and the severity of persistent or intense leukemia. A more uniform distribution of data points can be achieved by using histogram levelling and straight difference extending with morphological methods such as region opening, region closing, disintegration, and expansion. Thus, the Proposed Method is more effective than other previous methods.

Keywords--- Leukemia, Leukocytes, White Blood Corpuscles, Thresholding.

I. Introduction

Chronic lymphocytic leukemia (CLL) is a slow-growing disease. Many signs and symptoms do not manifest themselves until the disease has progressed to its terminal stage. At this point, medical science and doctors are unable to provide any further assistance to the patient. It has been established that chronic lymphocytic leukemia is associated with the development and spread of some of the conditions listed below: The presence of too few or too many red blood cells, which is the most common type of red blood cell, is another common cause of anemia. Anemia symptoms include weakness, fatigue, a lack of energy, and shortness of breath, which occur when abnormal white blood cells interfere with the ability of red blood cells to perform their functions properly. Leukopenia is a term that refers to both the production of antibodies by white blood cells and the ability of these cells to prevent disease. As a result of the abnormal white blood cells interfering with the ability of the average white blood cells to function properly, leukopenia develops. As a result of its use, it has the potential to cause immune system weakness, more frequent illness, and fever. One of the hallmarks of thrombocytopenia is the presence of a large number of clotting-related particles in the blood. As a result of this interference with the normal function of blood platelets, it is believed that gum disease, nosebleeds, and easy bruising have resulted as a result.

CLL is more common in the elderly, as evidenced by it affecting them at a higher rate than the younger population. According to the American Cancer Society, Chronic Myelogenous Leukemia can affect people of any age, but it is most common in people between the ages of 35 and 45. In addition to fatigue, weight loss, and pain in the left upper abdomen, Chronic Myeloid Leukemia (CML) is characterised by symptoms that appear at the time of diagnosis. A representation of CML cells in the form of a diagram is shown on the screen. Chronic myeloid leukemia is a type of cancer that begins in the white blood cells of the bone marrow and spreads throughout the patient's entire body.

II. Literature Survey

In order to diagnose specific diseases such as acquired immune deficiency syndrome, leukaemia, and other blood-related diseases, pathologists typically use an optical microscope to perform white blood cell (WBC) recognition as a first step, as described in [1]. Therefore, it is an extremely time-consuming, tedious, and expensive process that necessitates the expertise of subject-matter experts in the field who are knowledgeable about the subject matter. In this particular instance, a computer-aided diagnosis system that assists pathologists in their diagnostic work is clearly advantageous. When developing a computer-aided diagnosis system, one of the first steps is to separate white blood cells into different types of cells (WBCs). In this paper, white blood cells (WBCs) are

DOI:10.9756/INTJECSE/V14I5.475 ISSN: 1308-5581 Vol 14, Issue 05 2022

segmented from microscopic images using a computer program. To segment WBCs from microscopic images and extract nuclei from cell images, we use a three-stage approach that combines thresholding, k-means clustering, and modified watershed algorithms. We then separate cells and nuclei that overlap with one another. According to the results of the evaluation, the similarity measures, precision, and accuracy of the proposed method were 92.07 percent, 96.07 percent, and 94.30 percent, respectively, for nucleus and cell segmentation. The results of the manual segmentation and the results obtained through the proposed method are also strikingly similar, according to statistical analysis. The ability to adapt, to be dependable, and to be usable are all important characteristics for the next generation of health-related tools, as described in [2]. In addition to the fact that blood cell counting can be used to diagnose a wide range of diseases, it is also an important area of biomedical engineering research. In the present day, computers are used to determine both the red blood cell count and the white blood cell count in the human body, respectively (WBC). Methods requiring blood preparation are no longer recommended, but they may still be used under certain conditions or for diagnostic confirmation purposes only, even though they are no longer recommended. Recently, the MATLAB based computer-aided design software was used to develop algorithms that simplify the process of counting cells in a spreadsheet. Our algorithm, which has a computational performance of 1.98 seconds and a high level of digital count reliability, reduces the time it takes to complete an operation by 34% when compared to previous works. The Watershed transformation made it possible for these outcomes to be achieved. Moreover, the report's recommendations for healthcare delivery state that more severe pathology can be saved in digital files for future consultations, implying that hospitals will no longer be required. Using current image processing and segmentation techniques, we expect to develop a red and white blood cell algorithm to aid medical researchers. This method uses MATLAB to identify and locate infected White Blood Cells in the body.

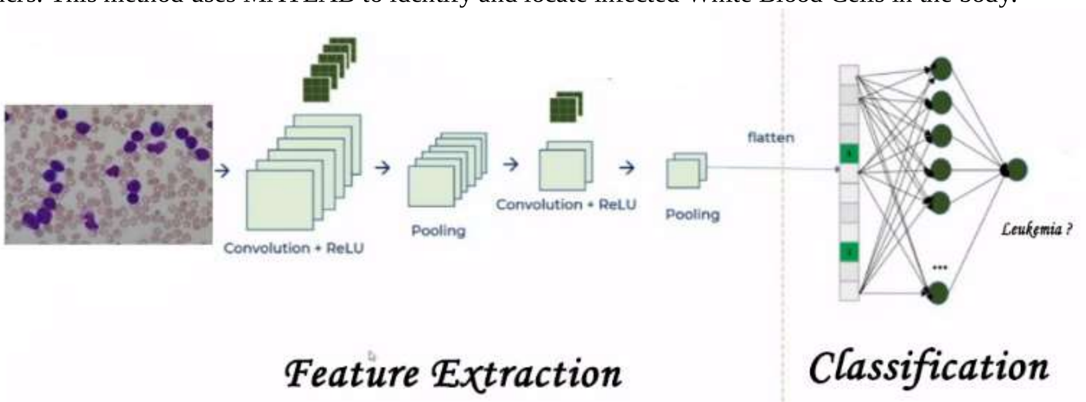


Figure 1: Proposed Framework for Leukemia Detection

Figure 1 shows the method of feature extraction and classification of CNN based leukemia detection. The classification of blood cells, according to [3] is more important than any other factor in the medical diagnostic system when it comes to analyzing and detecting disease. Once the different types of blood cells have been identified, the classification of blood-related diseases can be completed. Blood cancer known as leukaemia begins in the bone marrow and spreads throughout the body. The sooner you seek medical attention, the better. The sooner you get help, the better. The authors have devised a new genetic KNN pre-processing method for leukaemia images, which we present here. Genetic algorithm and KNN algorithm are used to remove noise from leukaemia image datasets in this paper. When working with a Leukemia image dataset, the median filter approach to image denoising and the G-KNN image enhancement are the two primary processes that make up the G-KNN pre-processing framework[4]. The proposed G-KNN method, which combines the K- and Genetic Algorithms, is intended to improve lung cancer detection classification accuracy.

By computerized analysis, white platelet tumor diseases such as leukemia and multiple myeloma are described in [5] as a testing biomedical investigation point for cancer that can be used to screen for cancer. Health problems must be identified and treated as soon as possible if they are to be prevented from becoming more serious in the future. If you are looking for cancerous growths, smear images are an effective and reliable source of information to use in your search[6]. When evaluating cancer growth discovery, such as a leukemia cancer image, it is critical to perform feature extraction and reduction in an efficient and effective manner. Barrow biopsy, lymph node biopsy, and lumbar puncture are some of the more common manual methods for diagnosing blood cancer that are currently available. Misdiagnosis and long computation times caused by manual detection of white blood cancer cells can result in death in some patients, which is a risk[7]. The use of automated detection methods, which have a lower rate of misdiagnosis than manual detection methods, can help doctors diagnose diseases more quickly and accurately. Specifically for this paper, we conducted a brief review of the current algorithms and methodologies used to detect

DOI:10.9756/INTJECSE/V14I5.475 ISSN: 1308-5581 Vol 14, Issue 05 2022

blood cancers, which you can read about here. We hope to accomplish this by providing a brief overview and review of related work based on artificial intelligence (AI), image processing, and machine learning techniques for automatically detecting blood cancer cells. Through the use of image processing techniques, it is possible to accelerate the detection of leukaemia in patients[8]. It is possible to overcome the subjectivity of the administrator's magnifying lens when the level of involvement and well-being of the spectator has an impact on the investigation. Image processing strategies can be used to overcome this influence. In order to detect cancer cells in white blood cells, automated methods must overcome a number of challenges, some of which are discussed in greater detail below. As soon as it is given the opportunity, automation should be able to correctly identify the intended goal. Support vector machines (SVMs) and k-means segmentation algorithms are two of the most widely used approaches for categorizing and segmenting data, respectively. The accuracy of an automated system is influenced by a variety of factors, including segmentation algorithms, feature extraction algorithms, and classification methods. Before feeding the features of the image into the classifier, it is necessary to extract the features of the image.

Researchers have been able to identify acute lymphocyte leukemia (ALL) cells in [9] cells using an efficient image processing algorithm. All types of cancer cells are more common in children, and if they are not treated early, they may succumb to the disease. The analysis is performed using the SVM (Support Vector Machine) algorithm. The Wavelet Transform algorithm was used to preprocess the image before it was displayed. The Confusion Matrix was employed to conduct a statistical analysis of the findings. According to the findings of this study, the success rate was 93.92 percent. After the morphological preprocessing was completed, image processing algorithms were constructed using a support vector machine (SVM), and the most effective iteration was selected from among the various iterations[10].

The iteration values are achieved at the start of our project's development. The states that exist before and after the Wavelet Transformation are distinct from one another[11]. Water can be discovered by reducing and multiplying test and training data sets. When the amount of test and training data is reduced, the outcome changes, whereas when the amount of training data is increased, the outcome changes[12]. When the tenth address is called, there are 260 data nodes, 120 adjective test data, and 140 adjective data in this iteration, for a total of 140 adjective data. According to the results, a total of 93.92 percent of network performance is achieved when the program is in use. This follows the research findings. Increasing the contrast between the grayscale images will improve their overall appearance. The Otsu method is used to accomplish global thresholding, which is then followed by feature extraction. According to the extracted feature, a classifier determines whether the cells are affected or normal, and it does so with a % accuracy rate. It is necessary to use the kNN classification.

III. Proposed System

The original image is in the RGB colour space. When viewed in the yellow component of the RGB image, the leukocytes are completely black because they lack any yellow component. Contrast Stretching is used to increase the image contrast in order to more easily threshold the leukocytes. The Zack algorithm divides data into sections based on thresholds. Drawing a straight line between the highest and lowest histogram values (h[bmax]) is accomplished by this method (h[bmin]). The maximum and minimum values of the histogram are found using the triangle method equation. The letters bmax and bmin stand for the maximum and minimum grey levels, respectively, in grey scale. After that, we will calculate how far the marked line is away from the histogram values between min and max, which will be the next step. This is the last and most important step. We can calculate the threshold value by increasing the d-distance while simultaneously increasing the intensity. Use other software to remove everything except your leukocytes from the image to achieve this effect. This will remove the background from the image. The Zack Algorithm is used to determine the appropriate threshold for background removal.

Image analysis becomes significantly more challenging to perform when leukocyte agglomerates are visible in blood images. The previous phase produced an image consisting solely of white blood cells; therefore, this is the only time we will be able to detect and separate leukocyte agglomerates during this phase (WBCs). The watershed segmentation technique is used to separate the leukocytes after being grouped.

DOI:10.9756/INTJECSE/V14I5.475 ISSN: 1308-5581 Vol 14, Issue 05 2022

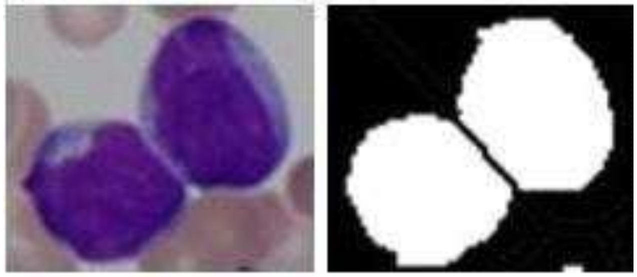


Fig. 2: Image before and after Applying

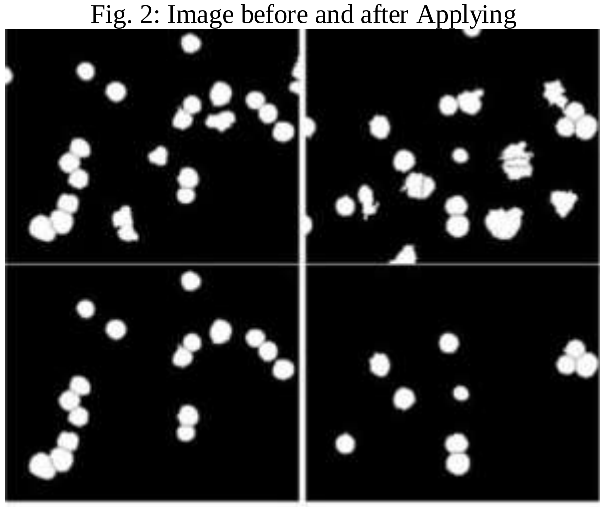


Fig. 3: Final Separation Results and Image Cleaning Results

Images must be cleaned to prevent errors from occurring during the subsequent stages of the analysis. Removing all abnormal components and leukocytes that have fallen off the image's edge is essential for accurate analysis (non-leukocytes). Using this tool, you can carry out simple operations, such as cleaning the image's outer edge. Each leukocyte's area and convex area are calculated separately at the outset. When identifying and eliminating components with irregular dimensions, it is necessary to compute the area of the area's mean of the component. As the images in the database contained an arrow, a Matlab script was written to remove the arrow from all images. The following are examples of images taken before and after the operation.

DOI:10.9756/INTJECSE/V14I5.475 ISSN: 1308-5581 Vol 14, Issue 05 2022

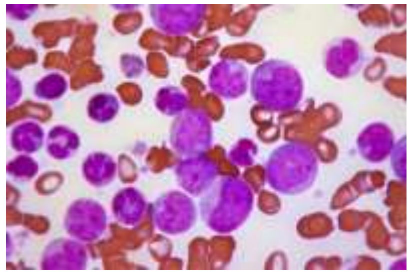


Fig. 4: Original Image

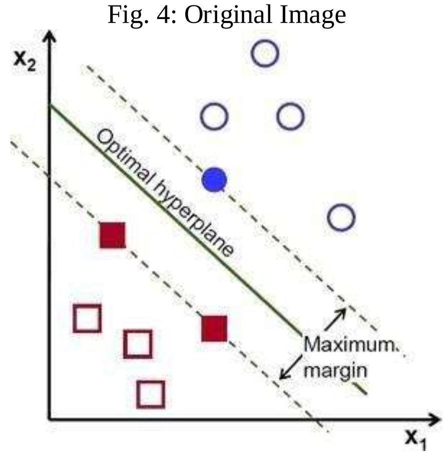


Fig. 5: SVM

The classification process was carried out using 3 features (Roundness, Elongation and Eccentricity) using the Support Vector Machine (SVM) classier because this model is particularly suitable for binary classification problems for which the separation between classes depends on a large number of variables.

IV. Simulation Results

The GLDM and GLCM are used to extract the features of the model. Cancerous microscopic images are classified using the SVM classifier.

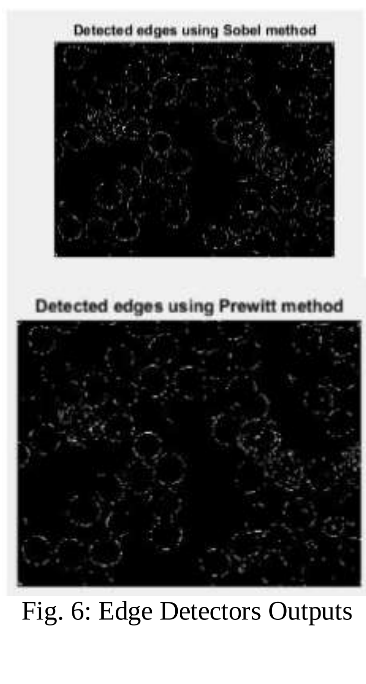


Fig. 6: Edge Detectors Outputs

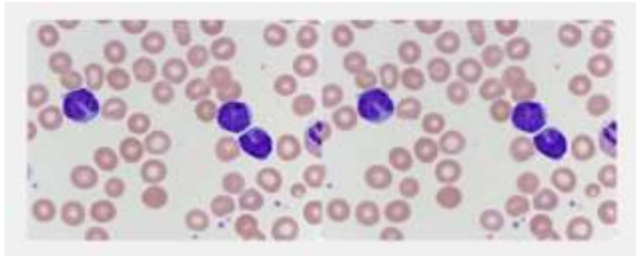
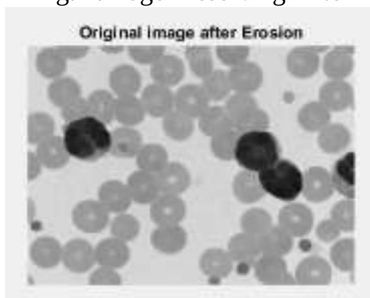


Fig. 7: Edge Preserving Filter



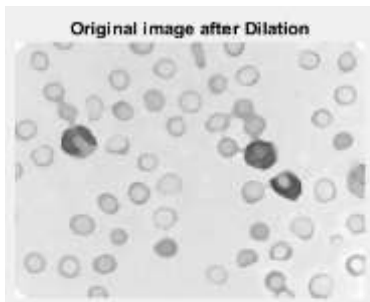
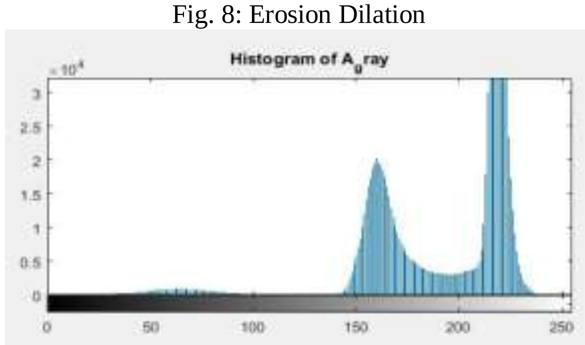


Fig. 8: Erosion Dilation



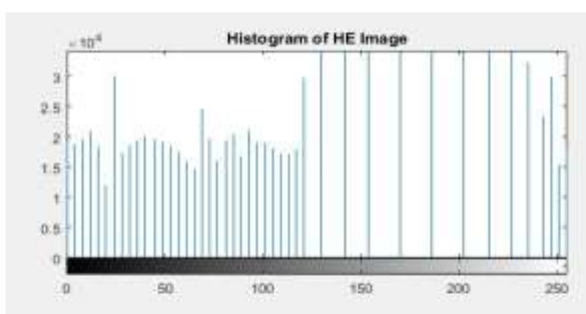


Fig. 9: Histogram Images

DOI:10.9756/INTJECSE/V14I5.475 ISSN: 1308-5581 Vol 14, Issue 05 2022

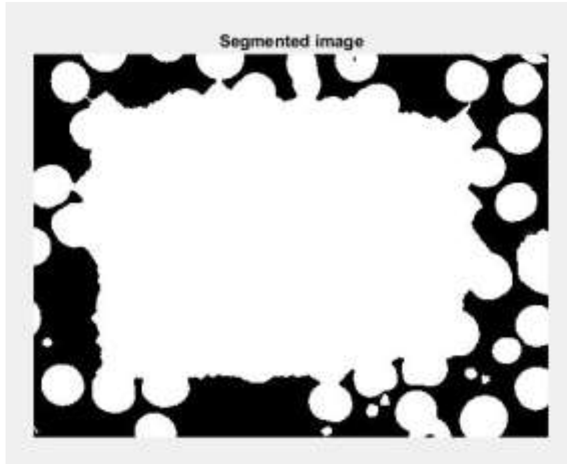


Fig. 10: Noise Reduction

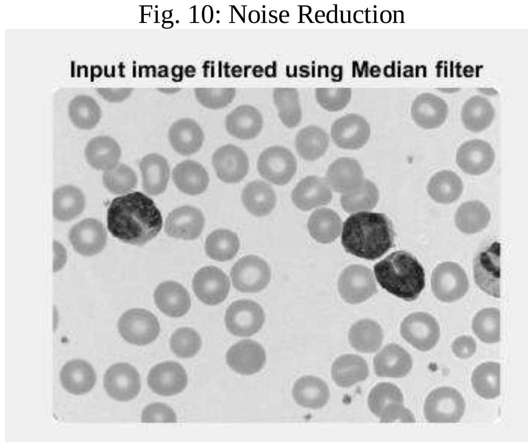


Fig. 11: Object Segmentation

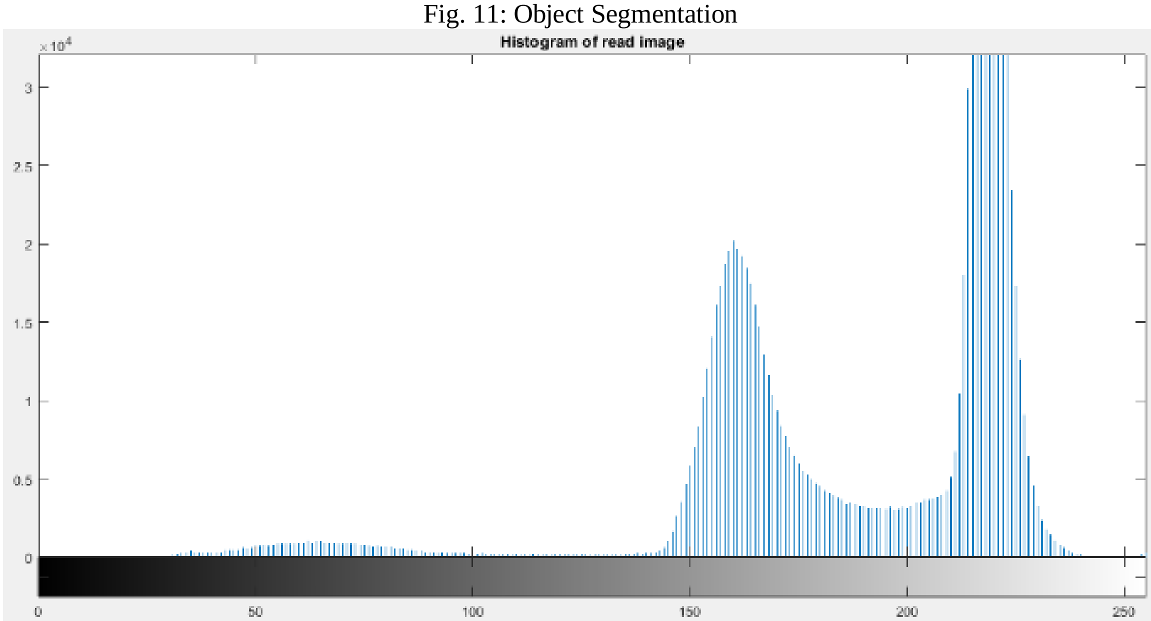


Fig. 12: Histogram Single Channel

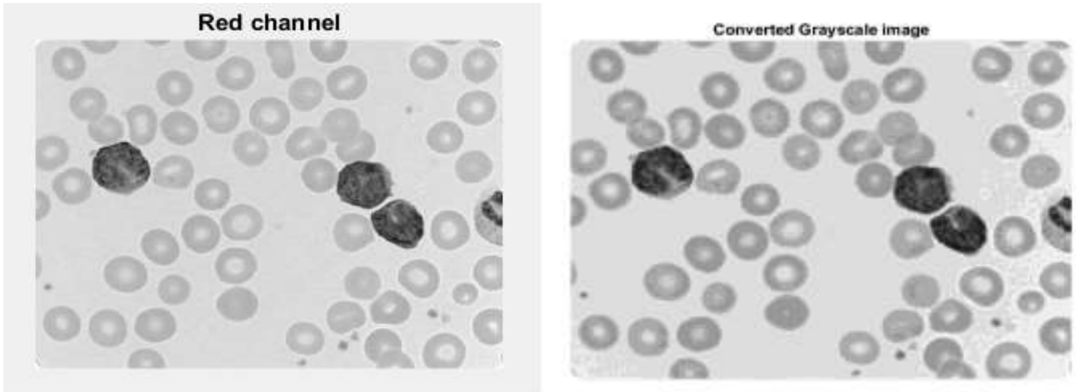


Fig. 13: RGB to GRAY

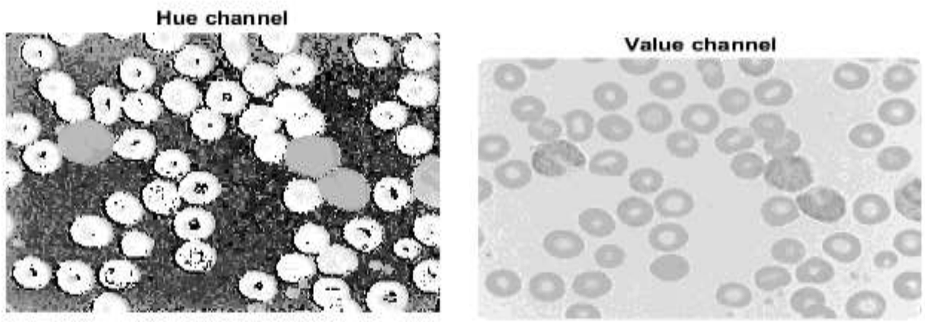


Fig. 14: RGB to HSV

The above-mentioned figures show the low-cost detection method of Leukemia blood cancer.

V. Conclusion

This research aims to develop a small image processing method that can distinguish between red blood cells and young white cells to detect leukemia. In today's world, the only way to detect blood clots is by visualizing minuscule photographs, including looking for surface irregularities, calculation, shading, and quantitative examination. Leukemia is one of the top three causes of death in humans. Early detection and identification of infection are critical for determining success and prognosis. The researchers hope to use this project to identify and rigorously test leukemia-affected cells to understand the disease better. The presence of immature cells and the persistence or severity of leukemia can differentiate between the two types of leukemia. When immature cells differentiate, various methods are used to accomplish this. These include histogram leveling, straight difference expanding, and morphological techniques such as region opening and closing, disintegration, and expansion. Research has shown that the Proposed Method outperforms current procedures.

References

[1] Hsu, Ching-Hsien, Xing Chen, Weiwei Lin, Chuntao Jiang, Youhong Zhang, Zhifeng Hao, and Yeh-Ching Chung. “Effective multiple cancer disease diagnosis frameworks for improved healthcare using machine learning.” Measurement 175 (2021): 109145.

[2] Khatib, Omar, Simiao Ren, Jordan Malof, and Willie J. Padilla. “Deep learning the electromagnetic properties of metamaterials—a comprehensive review.” Advanced Functional Materials 31, no. 31 (2021): 2101748.

[3] Surya Sashank, Gundepudi V., Charu Jain, and N. Venkateswaran. “Detection of Acute Lymphoblastic Leukemia by Utilizing Deep Learning Methods.” In Machine Vision and Augmented Intelligence—Theory and Applications, pp. 453-467. Springer, Singapore, 2021.

[4] Karami, Keyvan, Mahboubeh Akbari, Mohammad-Taher Moradi, Bijan Soleymani, and Hossein Fallahi. "Survival prognostic factors in patients with acute myeloid leukemia using machine learning techniques." PloS one 16, no. 7 (2021): e0254976.

[5] Tarek, Shahd, Hala M. Ebied, Aboul Ella Hassanien, and Mohamed F. Tolba. "White blood cells segmentation and classification using swarm optimization algorithms and multilayer perceptron." International Journal of Sociotechnology and Knowledge Development (IJSKD) 13, no. 2 (2021): 16-30.

[6] Mirmohammadi, Pouria, Marjan Ameri, and Ahmad Shalbaf. "Recognition of acute lymphoblastic leukemia and lymphocytes cell subtypes in microscopic images using random forest classifier." Physical and Engineering Sciences in Medicine 44, no. 2 (2021): 433-441.

[7] Mohammed, Zhana Fidakar, and Alan Anwer Abdulla. "An efficient CAD system for ALL cell identification from microscopic blood images." Multimedia Tools and Applications 80, no. 4 (2021): 6355-6368.

[8] Pavithra, M., R. Rajmohan, T. Ananth Kumar, and S. G. Sandhya. "An Overview of Convolutional Neural Network Architecture and Its Variants in Medical Diagnostics of Cancer and Covid-19." Handbook of Deep Learning in Biomedical Engineering and Health Informatics (2021): 25-49.

[9] Shehab, Lamia H., Omar M. Fahmy, Safa M. Gasser, and Mohamed S. El-Mahallawy. "An efficient brain tumor image segmentation based on deep residual networks (ResNets)." Journal of King Saud University-Engineering Sciences 33, no. 6 (2021): 404-412.

[10] Zhao, Gangming, Weifeng Ge, and Yizhou Yu. "GraphFPN: Graph Feature Pyramid Network for Object Detection." In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 2763-2772. 2021.

[11] Liu, Nanqing, Turgay Celik, and Heng-Chao Li. "Gated ladder-shaped feature pyramid network for object detection in optical remote sensing images." IEEE Geoscience and Remote Sensing Letters 19 (2021): 1-5.

[12] Quang, Tran Ngoc, Seunghyun Lee, and Byung Cheol Song. "Object detection using improved bi-directional feature pyramid network." Electronics 10, no. 6 (2021): 746.