Al-Salam Journal for Medical Science
Journal Homepage: http://journal.alsalam.edu.iq/index.php/ajbms
E-ISSN: 2959-53980, P-ISSN: 2958-0870

Leukemia detection using Artificial Neural Networks in Images of Human Blood Sample

Hakar J. Mohamed Salih[1], Jahwar Y. Arif[1], Shaimaa Q. Sabri[1], Ghada A. Taqa[2]*, Ahmet Çınar[3]

[1]Department of Computer Sciences, College of Science, University of Zakho, Duhok, Iraq.
[2]Department of Dental Basic Sciences, College of Dentistry, University of Mosul, Mosul, Iraq.
[3]Department of Computer Engineering, College of Engineering, Fırat, University, Elazığ, Türkiye.

*Corresponding Author: Ghada A. Taqa

DOI: https://doi.org/10.55145/ajbms.2024.03.02.01
Received November 2023; Accepted January 2024; Available online Febuary 2024

ABSTRACT: This article presents a preliminary report that uses minuscule images of blood tests to develop a diagnosis of leukemia. Examining through images is crucial since illnesses can be recognized and examined at an earlier stage using the images. The framework will be centered on leukemia and white blood cell illness. In fact, even the hematologist has trouble organizing the leukemic cells, and manually arranging the platelets takes a long time and is quite loose. In this way, early detection of leukemia recurrence allows the patient to receive the appropriate treatment. In order to address this problem, the framework will make use of the capabilities in small images and examine surface, geometry, shading, and quantifiable investigation modifications. These features' variations will be utilized as the classifier input. has transformed the use of images K proposes that (NN) and agglomeration. Examining a wide range of failure measures and increasing the intricacy of every system, the findings are examined. Utilizing feedforward (NN), image division is accomplished with less noise and a very sluggish conjunction rate. K-means agglomeration and (ANN) are intentionally used in this analysis to create a collection of processes that will work together to produce a much better presentation in (IS). An analysis has been conducted to determine the best rule for (IS).

Keywords: White Blood Cells, Cancer, Artificial neural, Read blood cell

1. INTRODUCTION

Medical imaging has played a crucial role in the development of important conceptual and interpretive techniques in environmental science and medicine over the past decade [1]. During this time, there was a notable progress in the recognition, recording, sharing, analysis, and presentation of healing images. As a result, the use of computerized image processing systems for therapeutic reasons has advanced significantly [2]. The hardest part of creating restorative images is coming up with integrated frameworks for the clinical sector. Close multidisciplinary collaboration between doctors and architects is essential for the development, execution, and endorsement of intricate treatment regimens. Hence, it is imperative to utilize a technique that can effectively and expeditiously distinguish between different categories of platelets during an urgent situation [3].[4].

Currently, the method used to distinguish blood types is visual examination of extremely small platelet pictures. Based on observable signs of blood issues, it may cause the ordering of certain blood-related diseases. One of the most feared infections in humans is cancer [5]. One form of plasma disease that can be lethal if not identified in a timely manner is leukemia [6]. Leukemia develops when the bone marrow generates an excessive amount of irregular white blood cells (WB) [7].

Grouped based on the type of WB cell that went through L remodeling. When a cancerous neoplasm is very far along, the abnormal blood cells are usually young effects, or young cells that are not working right. These cells multiply swiftly. Severe leukemia will rapidly deteriorate if treatment is delayed. Continuous leukemia allows access to young platelets, but as the illness worsens, beneficial cells are also generated. The symptoms of chronic leukemia develop gradually, and the disease takes longer to get worse [8].

2. MATERIAL AND METHODS

2.1 DATASET

The dataset used in the study is public. This dataset, called Blood Cancer - Image Dataset, was taken https://www.kaggle.com/datasets/akhiljethwa/blood-cancer-image-dataset/data .This dataset contains 10,000 single-cell images (64x64 pixels) taken from peripheral blood smears of patients diagnosed with Leukemia (Blood Cancer)

2.2 MODUALS

There are four modules used in developed method:

1. IMAGE ACQUISITION: AML typically has a larger and more suitable architectural structure, with distinct nucleoli that are separated by the Auer rode. However, it contains only one nucleolus within each nucleolus. Leucocytes appear darker after inhalation, while red blood cells (RBC) are visible at a moderate intensity. Based on these results, white cells look darker in photos than red blood cells (RBCs), which look lighter. Platelets are significantly smaller in size compared to red and white blood cells [9].

2. IMAGE Pre-processing: Images are saved in JPEG format throughout the acquisition process to reduce the amount of computing resources required and to ensure that there is no over-division in the watershed calculations because an image will contain more subtleties. As is typically the case, all blood component hues in the obtained images are close to the foundation shading; red and white platelets are gathered together, and the presence of stain and clamor in the blood slides is important. By enlarging their difference, they institutionalized the picture in an attempt to survive or lessen the impact of such variables.

The photographs will be disturbed by noise in the middle of expensive recoloring and picture security. The noise could be caused by shadows or light that cause a ROI to appear as a blurry image location. Since white platelets will be the ROI, the foundation will not be allowed. In order to improve distinctiveness, a picture upgrade was carried out, which can improve the quality of a restored image. [10].

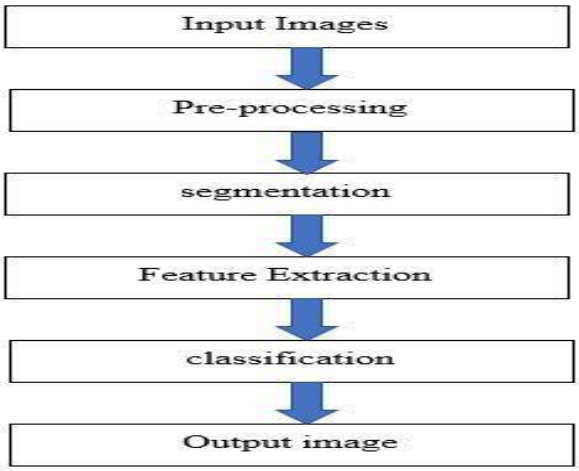

FIGURE 1. - Leukemia detection phases

3. IMAGE SEGMENTATIO: Three important steps are associated with sectioning the picture agreed:

1. Division of cellular element
2. Symmetry of stone and cytoplasm
3. Season of interfere (BC)

4. FUTURE EXTRACTION: The primary challenge in the era of platelet highlights lies in establishing a clear and precise definition that enables the accurate identification of diverse shot types. It is essential to eliminate geometric features from the core, such as area, length, boundary, balance, concave shape, adherence to rules, robustness, unpredictability, expansion, and structural element. Texture features include differentiation, entropy, connectedness, homogeneity, liveliness, and precise second energy. The approach involves utilizing the angle framework to analyze the image lattice histograms in the RGB, HSV, or L*a*b shading space, depending on the most appropriate option. Additionally, it considers color attributes such as mean shading values and statistical variables including mean value, variation, skewness, and kurtosis [11].

2.3 ALL, AML, CLL and CML

The task of assigning unknown test vectors to a designated lesson is known as rating. A support learning calculation is proposed in this development. The leukemia types will be categorized into ALL, AML, CLL, and CML using the RL method. The fundamental RL model is shown in Figure 2

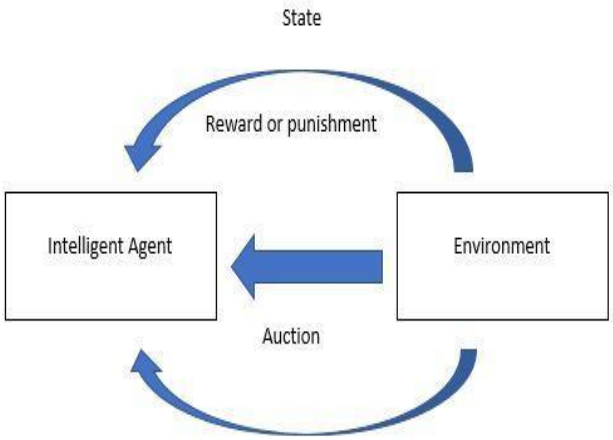

FIGURE 2.- Basic Model of RL

Although there are over a dozen rare forms of leukemia, four types primarily occur on a regular basis. These classifications are based on whether the leukemia is myelogenous versus lymphocytic and exceptional versus endless, that is: Granulocytic acute myelogenous leukemia (AML) Granulocytic Chronic Myelogenous Leukemia (CML). Given its widespread usage, Q-learning is the dominant technique in most applications of reinforcement learning. Q-learning is a sophisticated approach that assumes its Q-values were derived from the optimal Q-values, or Q(s,a), with little consideration for the specifics of the study. The Q-matrix will ensure the integrity of Q-characteristics. The conventional aggregate of forthcoming alterations, r, achieved by transferring, a, from the condition, s, is represented as Q(s,a). The movement with the highest vital Qvalue will be the most optimal. The Q(s,a) function will be activated or influenced by the experience in the following manner:

$$Q(s,a) = (1-\_)Q(s,a) + \_[r + \_ \max Q(s',a')]$$

where _ is the learning rate and 0< _ <1 is the refund factor. As an end, the examination theory can be viewed as in Figure 3.

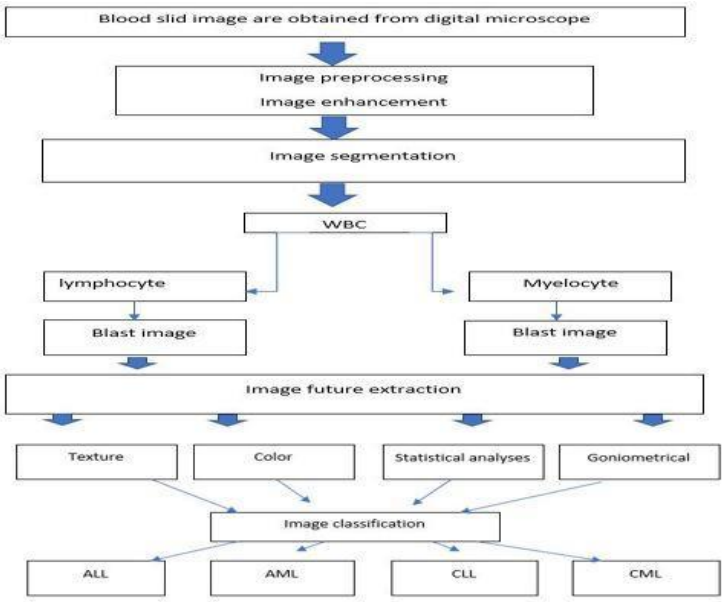

FIGURE 3. - The model algorithm work

2.4 CLASSIFICATION METHODS

In many fields, machine learning has become a vital tool for resolving challenging issues. Applying machine learning approaches to real-world problems has become simpler with the availability of machine learning tools like scikit-learn. Numerous categorization algorithms are available in the literature; we will list the ones we utilized in our investigation in the subsection that follows [12].

2.4.1 ARTIFICIAL NEURAL NETWORKS (ANN)

A machine learning model called an Artificial Neural Network (ANN) is modeled after the neural architecture of the human brain. It is made up of layers of connected nodes, or neurons. As data passes through these nodes, the weights of the connections are changed to identify patterns and generate predictions [13].

2.4.2 K-MEAN CLUSTERING ALGORITHM (K-NN)

K-NN is an incredibly straightforward machine learning method that aids in instance-based learning by helping us to answer categorization issues. The closest distance of the point from the training data, up to k data, is used to classify objects. All of the points that are accessible are stored, and additional points are categorized using similarity metrics (distance functions). The Euclidean distance between two points x1 and x2 is defined as follows:

$$Euclidean = \sqrt{\sum_{i=1}^{k} \left(x_1 - x_2\right)^2} \quad (1)$$

where k is the quantity of neighbors who are related. The ideal value of k is the one that minimizes the error, and we typically run the method multiple times with different values of k to determine the optimal values of k [14].

3. APPLICATION AND RESULTS

To implement the models, **MATLAB 2021 environment**. we have used on the windows 11 operating system installed on a computer with Intel(R) Core (TM) i5-8300H CPU @ 2.30GHz, 2304 MHz, 16 GB of RAM.

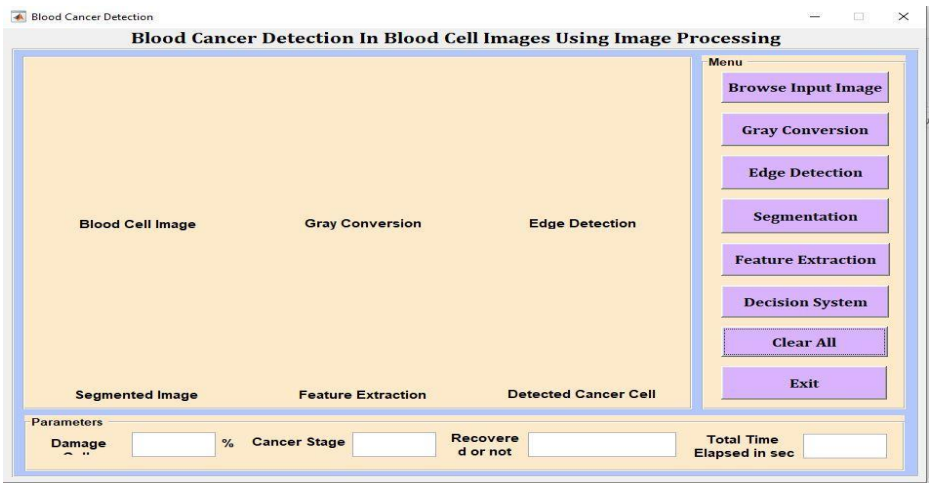

FIGURE 4.-Leukemia detection model

3.1 GRAY CONVERSION

In addition to the original image, this phase changes the image to grayscale. The first image is the enhanced version of the original, while the second is the image before the edge detection process's outcome.

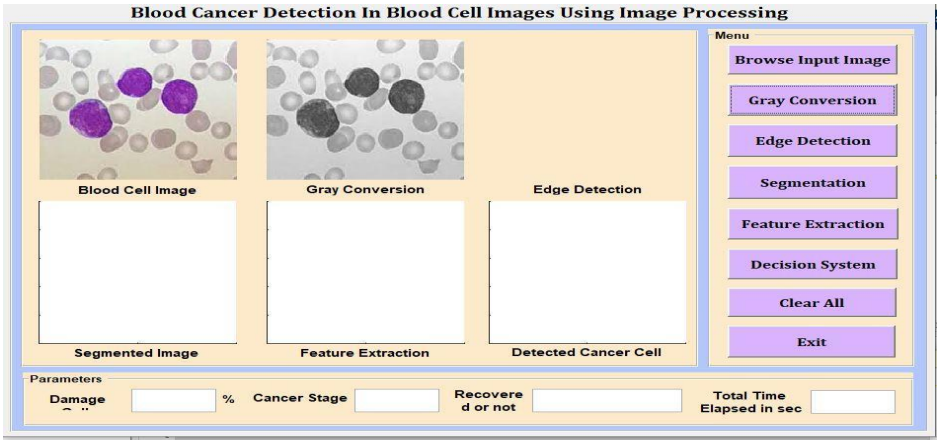

FIGURE 5.- Gray conversion

3.2. EDGE DETECTION

In various applications related to photo management, object recognition, and machine vision, edge detection plays a crucial role. The process of identifying and differentiating between acute discontinuities in advanced images, such as photometrical images, physical sections, and geometric locations, is known as edge finding. By limiting the pixels that display abrupt changes or discontinuities in a picture's power, edge detection aims to detect the presence of edges or boundaries in an image. It provides techniques for analyzing and limiting the optimal edge highlights to obtain important information for image analysis. Many related pixels that are located on the boundaries of a covering item and the image's foundation are called edges.

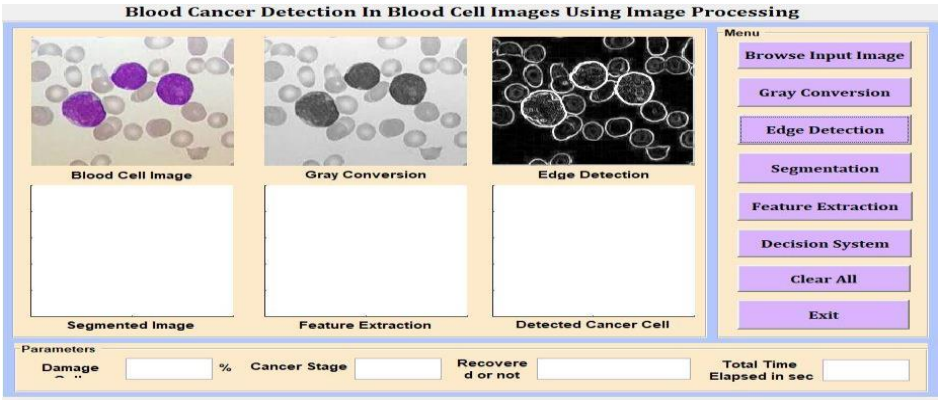

FIGURE 6. - Edge detection step

3.3 SEGMENTATION

The splitting of a picture into multiple pieces is the goal of the picture division. In this instance, the blood test image's picture division separates the white platelet (WBC) from the red blood cells (RBCs) and plasma. In the realm of image processing, the picture division process is a hot topic for research. Numerous studies have been conducted on the division of photos, and different division computations exist; nonetheless, no single calculation is suitable for all types of pictures. Accordingly, a computation meant for a particular type of image cannot be linked to a different type of image. As such, developing a single division technique for many images continues to be a work in progress. A preprogrammed image segmentation process that saves time and produces accurate results, especially for therapeutic images, can help save lives by detecting illnesses early and providing early treatment, as shown in figure 7.

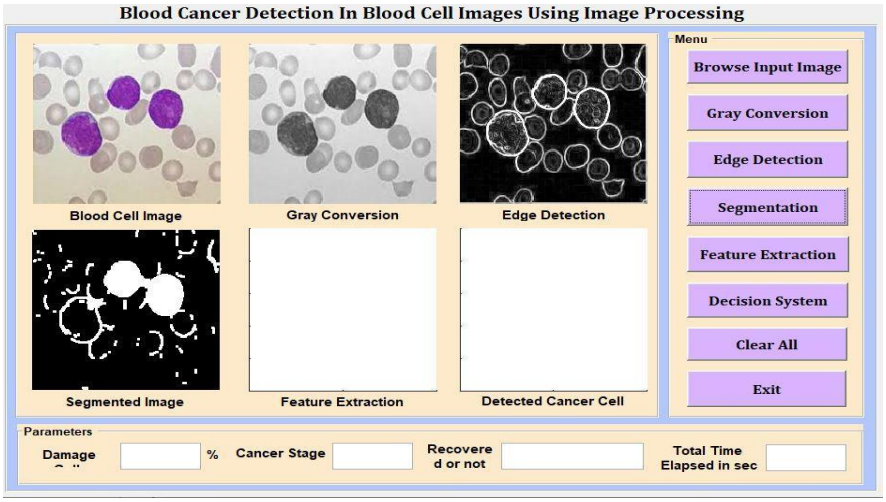

FIGURE 7. - Segmentation step

3.4 FEATURE EXTRACTION

The process of separating the desired highlights from the pre-processed images that have numerous anomalies is known as feature extraction. The image's standout features may include its estimation, shape, organization, area, and so forth.

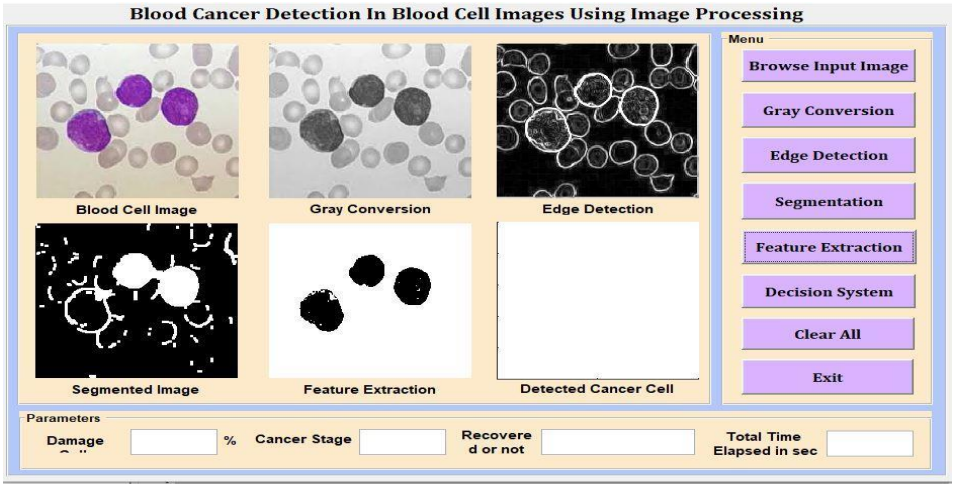

FIGURE 8. - Feature extraction step

3.5 EVALUTED THE RESULT

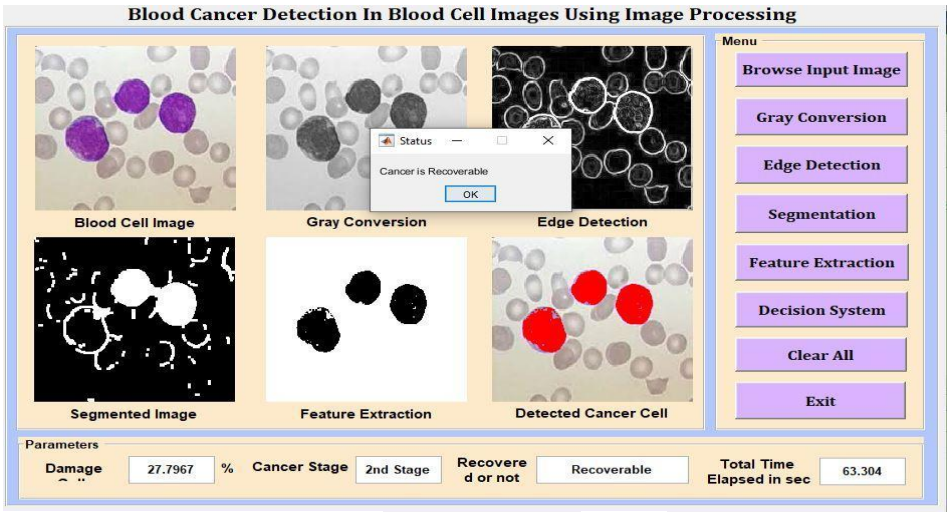

FIGURE 9.- Model process

By comparing our machine learning models, we assessed how well they could identify leukemia in human blood. However, because the dataset sizes utilized in the prior studies varied, it was difficult to compare their performance directly. During the first phase of the LEUKEMIA, many previous research had trouble getting enough samples, but over time, more samples were available via open sources like GitHub and Kaggle. Tables 1 provide the findings of our comparison investigation.

Table 1. - Parameter Output

Damage	Cancer stage	Recovered or not	Total time
33.83	2nd stage	Recoverable	63.06
29.33	2nd stage	Not Recoverable	63.11
23.33	2nd stage	Recoverable	65.5
25.21	2nd stage	Recoverable	61.45
27.56	2nd stage	Not Recoverable	61.35
34.11	2nd stage	Recoverable	56.15
21.46	2nd stage	Recoverable	67.6
25.56	2nd stage	Recoverable	61.33
33.78	2nd stage	Not Recoverable	60.06
27.36	2nd stage	Recoverable	58.36
23.33	2nd stage	Not Recoverable	64.56

Table displays parameter output damage. 33.83 percent of all image cancer cases are in the second stage, took 63.06 months to complete, and are curable. The second display damage, which is 23.33 cancer stage, cannot be recovered; 64.56 hours are needed in total.

4. CONCLUSION

In this test, microscopic blood test images are used to identify the various kinds of leukemia. The system will recognize surface, geometry, shadow, and actual inspection changes as a classification input based on the features in small photos. The company would be strong, interesting, less time-sensitive, more accurate, etc. The distribution of blood malignant stem cells and the classification of leukemia kinds using microscopic image processing are included in this investigation.

Funding

None

ACKNOWLEDGEMENTS

The authors would like to thank the reviewers for providing useful suggestions, allowing for the improved presentation of this paper.

CONFLICTS OF INTEREST

The authors declare no conflict of interest

REFERENCES

[1] S. Shafique and S. Tehsin, "Acute Lymphoblastic Leukemia Detection and Classification of Its Subtypes Using Pretrained Deep Convolutional Neural Networks," Technology in Cancer Research & Treatment, 2018. doi: 10.1177/1533033818802789.

[2] Q. Xiang, W. Xu, H. L. Wang, X. Y. Zhuang, W. C. Xie, and T. Rabczuk, "DigiSim—an open source software package for heterogeneous material modeling based on digital image processing," Advances in Engineering Software, vol. 148, p. 102836, 2020. doi: 10.1016/j.advengsoft.2020.102836.

[3] N. Ahmed, A. Yigit, Z. Isik, and A. Alpkocak, "Identification of leukemia subtypes from microscopic images using convolutional neural network," Diagnostics, vol. 9, no.3, p. 104, 2019. doi: 10.3390/diagnostics9030104.

[4] L. Boldú, A. Merino, S. Alférez, A. Molina, A. Acevedo, and J. Rodellar, "Automatic recognition of different types of acute leukaemia in peripheral blood by image analysis," Journal of Clinical Pathology, vol. 72, no. 11, pp. 755-761, 2019. doi: 10.1136/jclinpath-2019-205949.

[5] S. Mohapatra, D. Patra, and S. Satpathy, "An ensemble classifier system for early diagnosis of acute lymphoblastic leukemia in blood microscopic images," Neural Comput & Applic, vol. 24, pp. 1887-1904, 2014. doi: 10.1007/s00521-013-1438-3.

[6] L. Putzu, G. Caocci, and C. D. Ruberto, "Leucocyte classification for leukaemia detection using image processing techniques," Artificial Intelligence in Medicine, vol. 62, no. 3, pp. 179-191, 2014. doi: 10.1016/j.artmed.2014.09.002.

[7] S. Mohapatra, S. S. Samanta, D. Patra, and S. Satpathi, "Fuzzy Based Blood Image Segmentation for Automated Leukemia Detection," in Devices and Communications (ICDeCom), 2011, pp. 1-5. doi: 10.1109/ICDECOM.2011.5738491.

[8] M. Madhukar, S. Agaian, and A. T. Chronopoulos, "New Decision Support Tool for Acute Lymphoblastic Leukemia Classification," Image Processing: Algorithms and Systems X; Proc. SPIE 8295, 2012. doi: 10.1117/12.905969.

[9] Y. Duan et al., "Leukocyte classification based on spatial and spectral features of microscopic hyperspectral images," Optics & Laser Technology, vol. 112, pp. 530-538, 2019. doi: 10.1016/j.optlastec.2018.11.057.

[10] H. Kuresan, J. S. Satish, and N. Shanmugam, "Analysis of blood cancer using microscopic image processing," in Soft Computing and Signal Processing: Proceedings of 3rd ICSCSP 2020, vol. 2, Springer Singapore, 2022, pp. 403-415. doi: 10.1007/978-981-16-1249-7_38.

[11] M. T. Tsakok and F. V. Gleeson, "The chest radiograph in heart disease," Medicine, vol. 46, no. 8, pp. 453-457, Aug. 2018. doi: 10.1016/j.mpmed.2018.05.007.

[12] A. A. Soofi and A. Awan, "Classification Techniques in Machine Learning: Applications and Issues," J. Basic Appl. Sci., vol. 13, pp. 459-465, Jan. 2017. doi: 10.6000/1927-5129.2017.13.76.

[13] M. Hassan and N. Amiri, "Classification of Imbalanced Data of Diabetes Disease Using Machine Learning Algorithms," Oct. 2019.

[14] S. Tavakoli, A. Ghaffari, Z. M. Kouzehkanan, and R. Hosseini, "New segmentation and feature extraction algorithm for classification of white blood cells in peripheral smear images," Scientific Reports, vol. 11, no. 1, p. 19428, 2021. doi: 10.1038/s41598-021-98599-0.