International Transactions on Electrical Engineering and Computer Science
T. S. Latha, Vol: 1, No: 1, pp: 39-41, Sep 2022

Recognition of Blood Cancer Using Different Classification Techniques

T. Swarna Latha*

Abstract: Several image processing methods or applications have been established to get the requisite details from microscopic images for the lewd and cost-effective development of patient verdicts. Acute Lymphoblastic Leukemia (ALL) is a category of additional frequency in children. The word 'Acute' suggests that leukemia can develop rapidly and prime to fatal death in a few months if not managed. Because of its nonspecific existence, erroneous diagnosis refers to the effects and manifestations of everything. And though it is challenging for hematologists to identify leukemia cells, manual blood cell sorting is time-consuming and unreliable.
Consequently, early leukemia diagnosis helps present the patient with the necessary care. The mechanism recommends that individuals in blood picture the leucocytes from blood cells respond to this issue and select lymphocyte cells. The morphological catalog of certain cells is measured, and leukaemia's involvement is eventually classified. A literature review of several methods used to identify cancer cells has been carried out in this article.

Keywords: Blood cancer detection, C-Means clustering, K-Means clustering, Genetic algorithm

1. Introduction

Leukemia is a category of cancer that typically affects the body's blood-forming cells. This is a sort of cancerous type marked by the use of several irregular white blood cells in body. Within the bone marrow, leukemia instigates and then blows against several other human body organs [1]. Both children and adults can cultivate leukemia disorder. Nonetheless, in leukemia, the non-standard cell does not behave in the same manner as normal white blood cells.

Article History
Received: 13-06-2022; Revised: 27-08-2022;
Accepted: 14-09-2022

*Corresponding author: Department of Computer Science and Engineering, Narayana Engineering College, Nellore-524003, Andhra Pradesh, India.

E-Mail: proftswarnalatha@gmail.com

Ph: +91-8555921789

The leukemia cells would begin to grow as well as distribute, eventually stationing out all the regular blood cells. The effect is that it becomes impossible for the body to combat pathogens, carry oxygen and manage to bleed [2]. In additional forms of cancer, initial detection of leukemia be situated critical to improving the possibilities of a person's cure. Therefore, an accurate diagnosis may be achieved under which a human specialist could collect details after slur blood tests, count the number of erythrocytes in addition to the number and division of leukocytes. While including other objects that may indicate either a particular illness or personality disorder; nevertheless, a human expert's examination of smear blood is tired of a tired condition [3-5]. For this cause, automated blood cell breakup research has been increasing in recent years. In this phase, the image isolates a copy into shared parts or fragments involving image pixels with similar information attribute morals. It is an important problem because it is the first step in interpreting pictures, and every other step, such as extraction and identification of features, relies heavily on its effects [6]. A blood smear image of an affected person with leukaemia cancer is represented in Fig. 1.

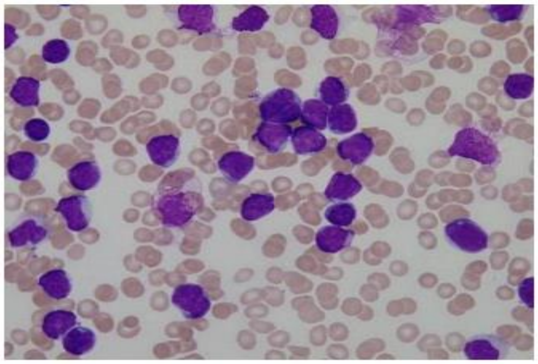

Fig. 1: Blood smear image of an affected person with leukaemia cancer

A correct subdivision is the most important phase in conducting a completely automated and strong vision arrangement for identifying and cataloging blood cells, particularly in medical image analysis, and several proposals exist to segment such pictures [7]. For example, a division technique based on k-mean, fuzzy Gustafson Kessel and Irregular Sets gathering with detection of the adjacent one division of the white blood cell nucleus, or flares, is conducted [8] consuming colour reproductions RGB and HSI, and later all subdivisions are compared according to the colour model. In an RGB image, the colour space is translated into HSI and the conduit S is segmented consuming a tentative beginning value. So called colour Structure-Code is hand-me-down to create an systematized graph, based on the assumption that images have homogeneous regions [9]. In several of the methods mentioned in the literature, a pre-processing stage must be enhanced before the division, thereby improving the processing period, even considering just division. On the other hand, after the initial evolutionary algorithms are anticipated in the 60's, many nature-inspired algorithms have appeared, making it evident that biological-inspired approaches can be effectively relocated into innovative computational standards [10].

2. Topical study works: A brief review

2.1 Canny based methods

This study's approach concerns the identification of the edges of leukemia in the picture of white blood cells. To this end, the obtained white blood cell picture is pre-existed for further study to eliminate the noise and towards smoother images. To diagnose the presence of leukemia cells, the next stage includes finalizing the blood [11-12]. Besides, edge detection techniques such as the Ant Colony Optimization (ACO) technique, Sobel, Prewitt and Robert gradient feature extraction techniques are used to remove the leukaemia cells' patterns from the blood image.

2.2 Image Fusion Methods (IFS)

In science, several fusion approaches are now possible for a day but any new approach is focused on the common features of the basic method [13-14]. This includes several fundamental strategies for image fusion. Individuals are IHS, PCA, BT, MRA, and EMD, respectively. Here are the logically mentioned strategies above.

A. Intensity-Hue-Saturation (IHS) Image Fusion Method

IHS is a popular way to combine high spatial resolution, single band, multispectral remote sensing signal, pan then low three-dimensional resolution. The multispectral copy's R, G and B bands are converted into HIS elements; swapping the strength portion with the pan image and doing the inverse transition to achieve a strong multispectral image spatial resolution (see Fig. 2). HIS will boost the multispectral image's spatial features and boost the textural properties of the fused picture. Still, the fusion image has a significant spectral distortion [15]. The transformation of the HIS is used for geological mapping since the transformation of the IHS will allow different types of spectral and spatial landscape information to be integrated for analysis into a single data collection.

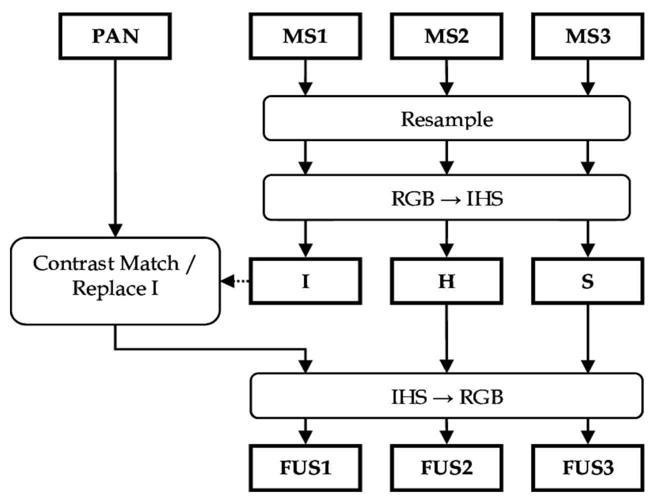

Fig. 2: Block diagram of IHS fusion method

While the HIS approach has been commonly used the method does not decompose a picture trendy frequency space, such as advanced or inferior frequency, into various frequencies. Therefore, the IHS approach should not be used to boost those picture characteristics. HIS technique's color distortion is also important. The PAN image is matched to the strength portion that extends before the reverse transformation to minimize the color distortion. Centred on the non-sub sampled Contour, image fusion enabled transformation and HIS improved preservation of spectral information and spatial details and better impact of integration. The segment-based fusion was precisely generated with HIS transformation or a spectral function preserving image merge coupled with spatial domain filtering.

B. Photo fusion of the brovey turn

BT is built on transition of chromaticity. It is an easy way to integrate data from different sensors with constraint that are only three levels elaborate. The goal is to regularize the trio multispectral groups used for RGB demonstration and multiply the outcome by any additional details needed to improve the component of strength or illumination of image. For the precise adaptation of criteria, this approach requires an experienced observer [16]. This creates the development of a user-friendly digital instrument. The Brovey Transform was developed to prevent the multiplicative process's drawbacks. A mixture of arithmetic operations normalizes the spectral bands until the panchromatic representation is multiplied.

C. PCA Infusion method

PCA conversion is a method for simplifying data collection from statistics. Pearson invented it in 1901 and Hotelling in 1933, though Jolliffe in 2002 is the strongest recent guide. The purpose of the approach is to decrease the dimensionality of multivariate data while retaining as many of the related data as possible [13-15]. The linked data collection is converted to an uncorrelated dataset. Mostly, PCA data is more interpretable than the source data. The usage of this approach may minimize the redundancy of the image data. The PCA of Fig.3, contains a statistical method that converts a set of associated variables into several key components called uncorrelated variables [17]. It computes the data set with a compact and optimal definition. The first primary element is taken with the highest variance along the road.

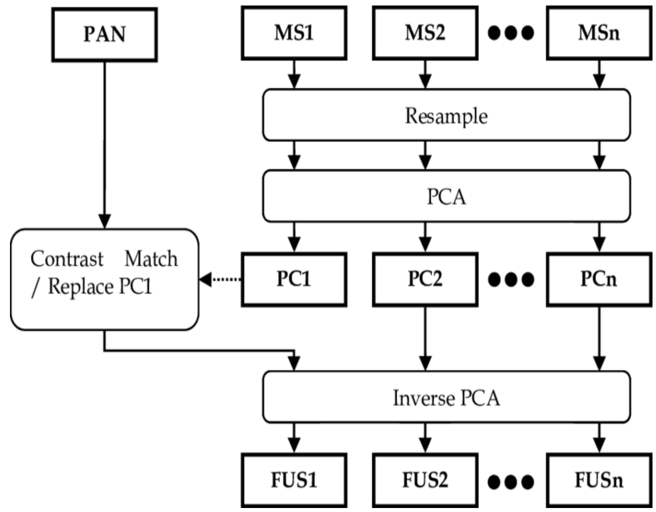

Fig. 3: Block diagram of a PCA fusion system

3. Proposed classification methods

3.1 Genomic algorithm

A populace of cords (called chromosomes or the genome genotype) in a genetic algorithm that encrypts intrant solutions (persons, animals, or phenotypes) for the optimization problem evolves towards improved resolutions. Resolutions are usually denoted as sequences of 0s to 1s in binary, although additional encodings also available. Generally, development begins since a group of people created spontaneously and occurs in centuries. Multiple entities are stochastically chosen from the existing population (Depending on their health) and adjusted (recombined and probably spontaneously mutated) to create a new population in respective generation, determining the suitability of individual people in populace. The current community is then included in the algorithm's next iteration [18]. The algorithm typically stops when also a cumulative number of generations have been situated generated or a suitable degree of suitability for the population has been situated attained as shown in Fig. 4. When owing to the maximum number of generations, the algorithm has stopped, a suitable answer may or may not have been found. Genetic algorithms are implemented in bioinformatics, computer science, architecture, economics, chemistry, logistics, mathematics, physics and other areas.

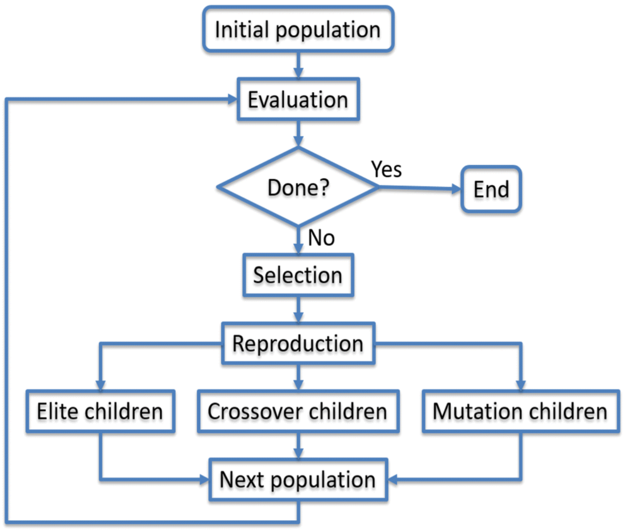

Fig.4: GA approach for leukemia image classification

The number of structures to contemplate is high, this heuristic solution has been selected [19]. The aim is to separate the maximum important relatives of characteristics before defining entities that consume the similarities measured according to these relatives. This algorithm's initial stage deals with the separation of the very less important structures from the wide collection. This won't specifically issue of traditional feature discovery recognized in Information Mining. We have principle is essential to choose less than 5 percent of the functions. But this issue is similar to the problem of classical feature selection, and as we noticed that most suited for glitches with a huge number of structures, we can use a genetic algorithm. There are various steps of the genetic algorithm considered here. For a immobile number of groups, it continues. The transformation is an operative that makes assortment. A chromosome has a risk of mutating at the mutation period. A genetic material chosen to transform, we casually pick many n bits to turn, so n bits are aimlessly selected and overturned. A probabilistic binary range of the contest is taken. Here a chromosome is a cord of bits whose scale equals the number of characteristics [20]. These operators allow the search space to be searched by GAs. Usually, though, operators have disruptive in addition to positive results. It is customized to the tricky. We use a Subgroup Size-Oriented General Function Cusp Operator (SSOCF), which preserves valuable knowledge slabs and generates the same offspring distribution as the parents, as shown in Fig.5. Offspring are kept only if they suit better than the population's least successful person. Off-springs maintain characteristics shared by the two parents and the non-transferable characteristics are acquired off-springs consistent to the ith parental with likelihood (ni - nc/nu) where ni is the number of nominated features of the ith parental, nc is the number of frequently picked characteristics from all pairing associates and nu is number of selected non-transferable characteristics.

The transformational operator that makes assortment. A chromosome has a risk of transforming at the mutation period. Chromosome is chosen to change, we aimlessly pick several n bits to turn, so n bits are aimlessly selected and overturned. A probabilistic binary range of the tournament is occupied. Collection of contests retains n contests to pick n people. Each contest contains of selection 2 population elements and selecting the better one with a [0.5, 1] likelihood. The Chromosomal Difference Produces a particular difference that is a gentle of bit-to-bit distance bit not available. I am taken into consideration, but the two people's whole span is compared. If in this window, one and only one person has a selected function, the distance is increased by one.

3.2 The K-mean clustering algorithm

K-mean is a technique of clustering annotations into an exact number of separate clusters. "K" corresponds to the number of specified clusters, as shown in Fig. 6. Different distance measurements occur to decide which reflection is to be added to which cluster. But this issue is similar to the problem of classical feature selection, and as we noticed that most suited for difficulties with a huge number of structures, we can use a genetic algorithm. There are various steps of the genetic algorithm considered here. For a secure number of generations, it continues [21]. Aiming algorithm is to reduce the measurement between the cluster's centroid and specified thought by adding an thought to each cluster iteratively and terminating the last distance measurement be situated reached.

Algorithm Summary

- Initially, the trial space is segregated into K bunches and the annotations are allocated to clusters randomly, Per specimen

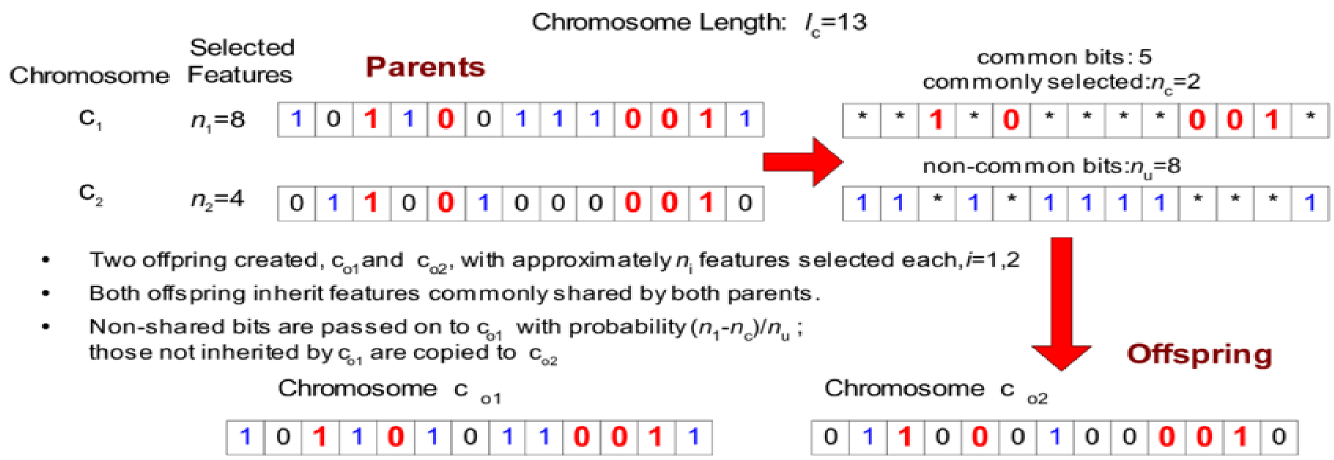

Fig. 5: SSOCF crossover operator

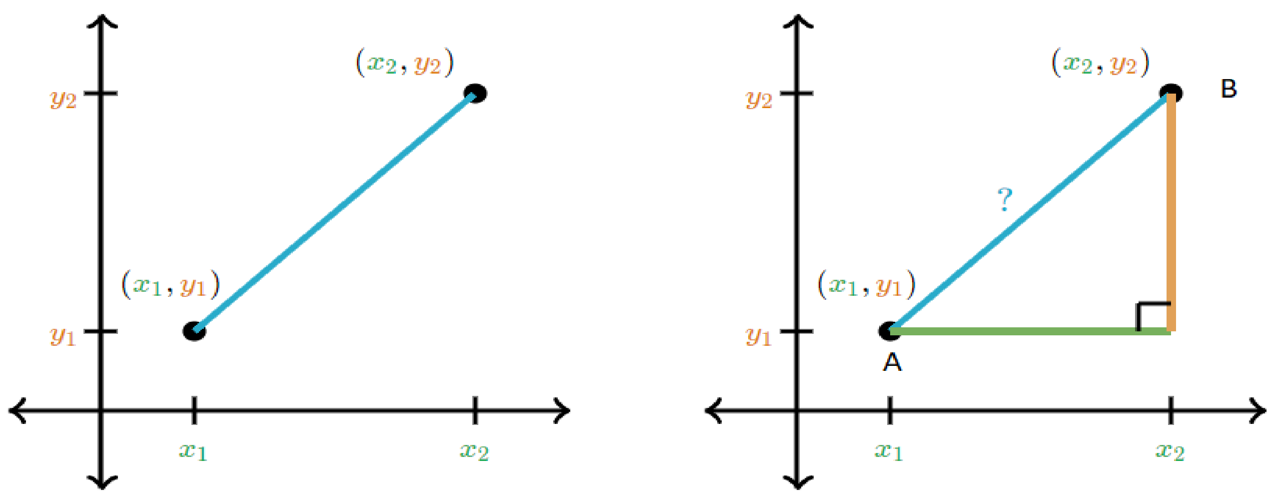

Fig. 7: Comparison between the Euclidean measure and Manhattan measure

- Calculate the distance of the cluster from reflection to the centroid.
- If the model is nearest to its bunch, then leave additional cluster chosen by ELSE.
- Repeat phases 1 and 2 before no findings from one cluster are transferred to another

The clusters are stable until phase 3 finishes and a cluster is allocated to each sample that leads to low management of available to both the centroid of the cluster

The Euclidean distance, the Euclidean square distance and the Manhattan or City distance are typical distance scales [22]. The thru geometric distance amid two points corresponds to the Euclidean scale. Using the squared Euclidean distance, this measures the directly above-squared length, i.e., a quicker way to estimate the reserve. The Manhattan metric measures a grid-based interval between points, as shown in Fig. 7.

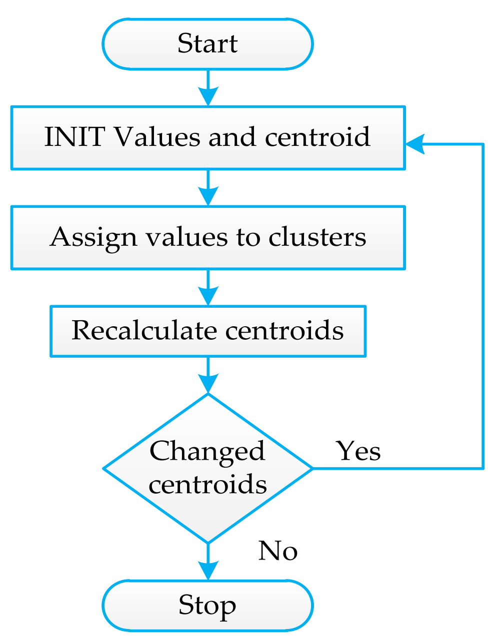

Fig. 6: Algorithm of K Means clustering

3.3 Fuzzy C-Means clustering

The method of grouping relevant data into groups or clusters is data plus clustering, such that objects in the same class are identical and items in separate classes are as different as possible [23]. The fuzzy C means algorithm is shown in Fig.8. Different similarity measurements can be used to position objects in groups based on the information's richness and the reason grouping is being used, where even the classification accuracy determines how the groups are created. Compared to hard segmentation, each stage has a degree of cluster membership in fuzzy clustering, as in fuzzy logic, instead of belonging exclusively to only one cluster [24].

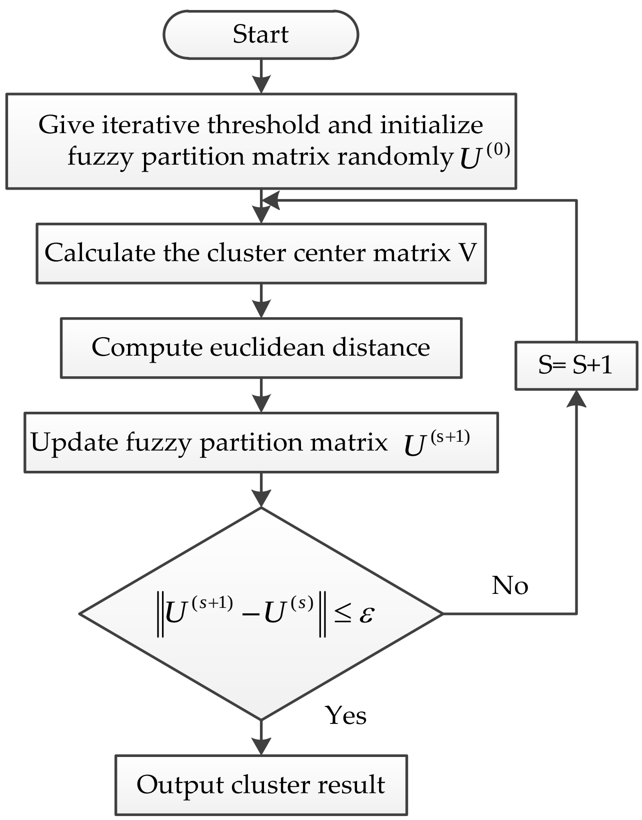

Fig. 8: Algorithm of a Fuzzy C- Means clustering

4. Results

In computer simulations are carried on the MATLAB environment for three different input images and carried the outputs for three input images with different sizes which can be found from Fig. 9 to Fig. 20. These output classifications in terms of PSNR, MSE value remain recorded popular Table.1. The PSNR value very high compared to the other two techniques for the GA approach and found better

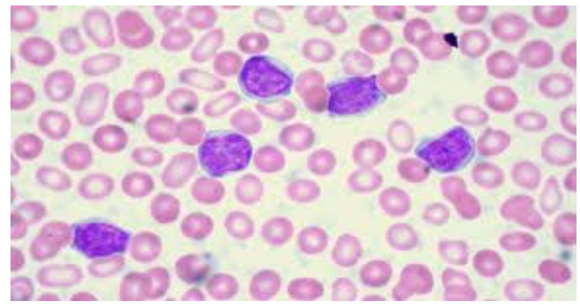

Fig. 9: Input picture 1

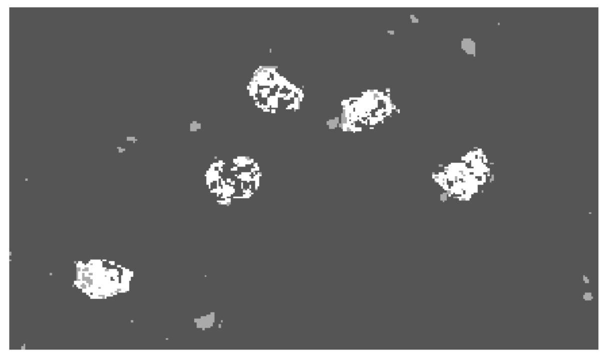

Fig. 10: Genetic Algorithm for input picture 1

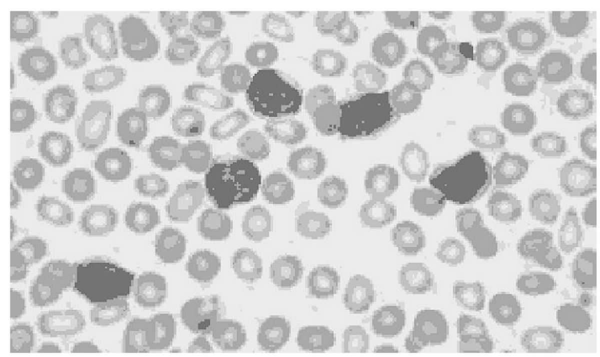

Fig. 11: K-mean clustering for input picture 1

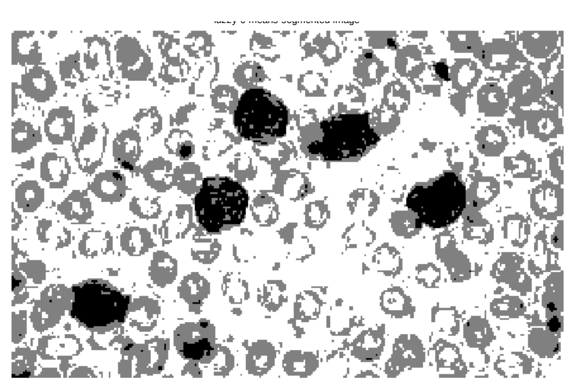

Fig. 12: C-mean clustering for input picture 1

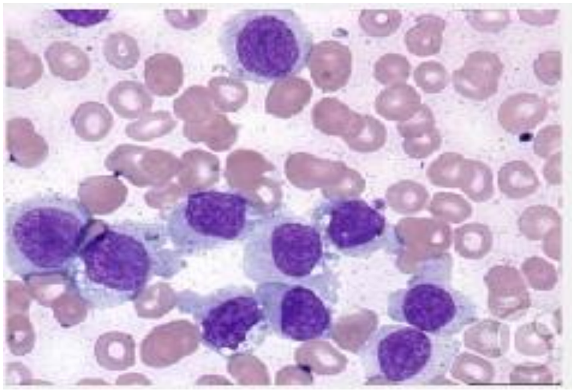

Fig. 13: Input Picture 2

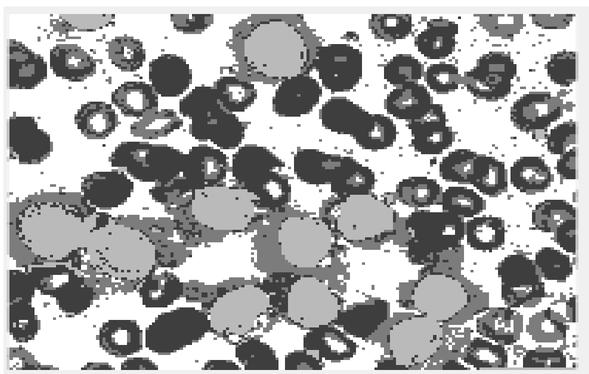

Fig. 14: Output of C-mean clustering for input image 2

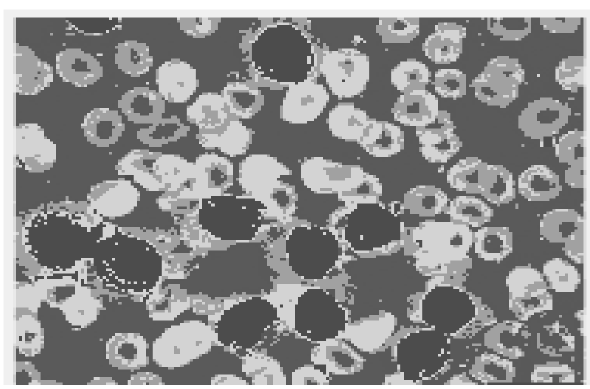

Fig. 15: Output of K-mean clustering for input image 2

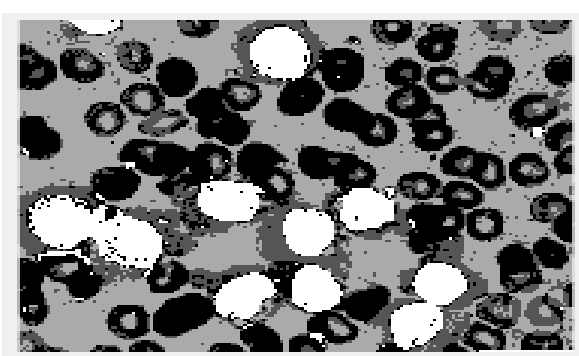

Fig. 16: Output of Genetic Algorithm for input image 2

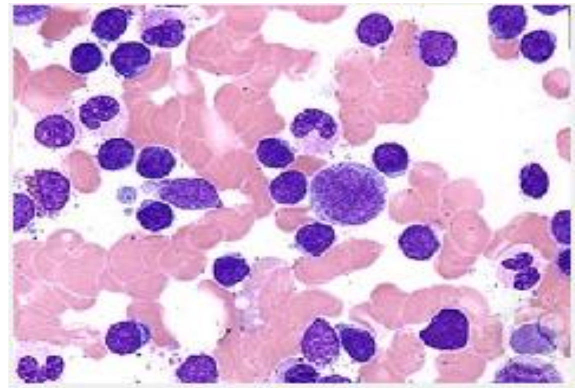

Fig. 17: Input picture 3

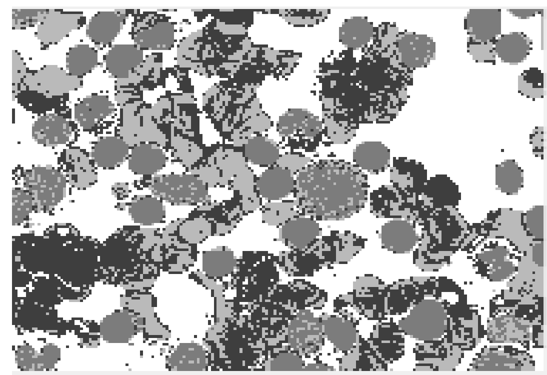

Fig. 18: Result of C-mean clustering for input image 3

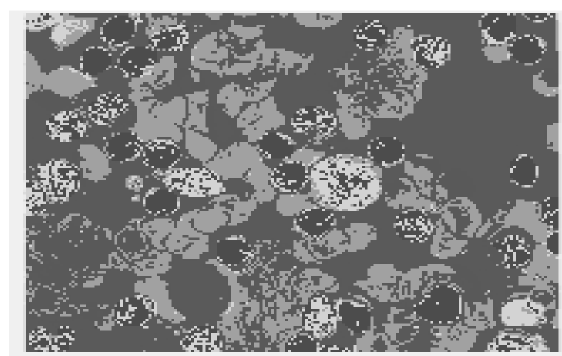

Fig. 19: Result of K-mean clustering for input image 3

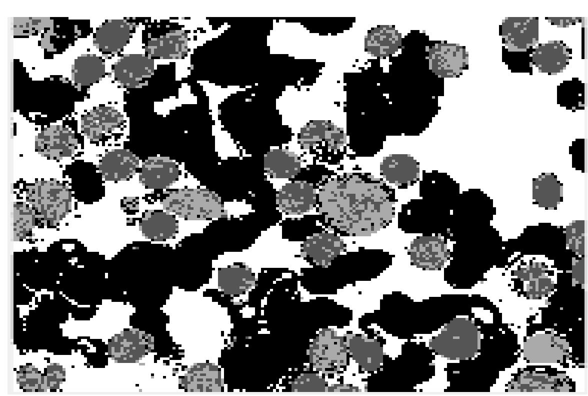

Fig. 20: Output of Genetic Algorithm for input image 3

Table. 1: Comparison between the K-mean, C-mean clustering and Genetic algorithm

Method/ Value	Input	PSNR Value	MSE Value
K- Means	Input image 1	1.6366	4.4609
C- Mean		1.8986	4.1997
GA		6.2213	1.5521
K- Means	Input image 2	1.76	5.3127
C- Mean		2.587	5.1926
GA		6.771	1.2376
K- Means	Input image 2	1.9366	5.6281
C- Mean		2.8986	5.2861
GA		6.937	1.4562

5. Conclusion

This study includes identifying leukemia forms using microscopic blood sample photographs. By utilizing characteristics in microscopic pictures, the device can be constructed by analyzing adjustments in texture, geometry, colors and statistics as a classification input. The device should be effective, accurate, lower processing time, lower error, high performance, cheaper expense, durable for individual varieties, sample selection protocols, time, etc. By rapidly detecting, addressing and managing blood disorders for a single patient, knowledge derived from microscopic photographs of blood samples will help people. It is found from the above analysis; the GA-based approach detects leukaemia well compared to K-means and C-means bunching.

Acknowledgment

The author is very thankful to the management of MREC (A), for their support to carry out this research work

Conflict of Interest

The author declares that they don't have any conflict of interest

References

[1] C. Raje, J. Rangole, "Detection of Leukemia in microscopic images using image processing," *International Conference on Communications and Signal Processing (ICCSP)*, pp. 255-259, April 2014.

[2] K. Deb, A. R. Reddy, "Reliable classification of two class cancer data using evolutionary algorithms", *Bio Systems*, Vol. 72, pp.111–129, 2003.

[3] S. Mohapatra, S. S. Samanta, D. Patra and S. Satpathi, "Fuzzy based Blood Image Segmentation for Automated Leukemia Detection", *International conference on devices and communications*, Mesra, pp. 1-5, 2011.

[4] N. Patel, A. Mehra, "Automated Detection of Leukimia using microscopic images", *Procedia computer science*, Vo. 58, pp. 635-642, 2015.

[5] I. Jafar, H. Ying, A. F. Shields, O. Muzik, "Computerized Detection of Lung Tumors in PET/CT Images", *International Conference of the IEEE Engineering in Medicine and Biology Society*, New York, pp. 2320-2323, 2006.

[6] M. Nakao, A. Kawashima, M. Kokubo, K. Minato "Simulating Lung Tumor Motion for Dynamic Tumor Tracking Irradiation", *Nuclear Science Symposium Conference Record*, Honolulu, pp. 4549-4551, 2007.

[7] E. Donald "Introduction to Data Mining for Medical Informatics", *Clinics in Laboratory Medicine*, Vol. 28, No. 1, pp. 9-35, 2008.

[8] R. Zhang, Y, Katta, "Medical Data Mining," *Data Mining and Knowledge Discovery*, pp. 305-308, 2002.

[9] I. M. Mullins "Data mining and clinical data repositories: Insights from patient data set," *Computers in Biology and Medicine*, vol. 36, pp. 1351- 1377, 2006.

[10] Zadeh, H. Ghayoumi, S. Janianpour, and J. Haddadnia, "Recognition and Classification of the Cancer Cells by Using Image Processing and Lab VIEW", *International Journal of Computer Theory and Engineering*, Vol. 5, No. 1, pp. 104-107, 2013.

[11] Z. Diana, Y. Pigovsky, and P. Bykovyy. "Canny-based method of image contour segmentation", *International Journal of Computing*, Vol. 15, No. 3 pp: 200-205, 2016.

[12] Biswas, Ranita, and Jaya Sil. "An improved canny edge detection algorithm based on type-2 fuzzy sets", *Procedia Technology*, Vol. 4, pp: 820-824, 2012.

[13] Z. Wang, D. Ziou, C. Armenakis, D. Li, & Q. Li "A comparative analysis of image fusion methods", *IEEE transactions on geosciences and remote sensing*, Vol. 43, No. 6, pp. 1391-1402, 2005.

[14] H. Ghassemian "A review of remote sensing image fusion methods", *Information Fusion*, Vol. 32, pp. 75-89, 2016.

[15] F. Vollnhals, J. N. Audinot, T. Wirtz, M. M. Bonin, I. Fourquaux, B. Schroeppel & S. Eswara "Correlative microscopy combining secondary ion

mass spectrometry and electron microscopy: comparison of intensity–hue–saturation and Laplacian pyramid methods for image fusion", *Analytical chemistry*, Vol. 89, No. 20, pp. 10702-10710, 2017.

[16] K. C. Bhataria and B. K. Shah, "A Review of Image Fusion Techniques," *International Conference on Computing Methodologies and Communication*, Erode, pp. 114-123, 2018.

[17] L. B. Kone, N. M. Kunda, T. B. Tran, & A. V. Maker "Surgeon-placed continuous wound infusion pain catheters markedly decrease narcotic use and improve outcomes after pancreatic tumor resection", *Annals of surgical oncology*, pp. 1-9, 2020.

[18] S. Belciug, & F. Gorunesc, "A hybrid neural network/genetic algorithm applied to breast cancer detection and recurrence", *Expert Systems*, Vol. 30, No. 3, pp. 243-254, 2013.

[19] P. D. Cesar, R. P. Ramos, and M. Z. D. Nascimento "Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm", *Computer methods and programs in biomedicine*, Vol. 114, No. 1, pp. 88-101, 2014.

[20] P. Bhuvaneswari & A. B. Therese "Detection of cancer in lung with K-NN classification using genetic algorithm", *Procedia Materials Science*, Vol. 10, pp. 433-440, 2014.

[21] B. Zheng, S. W. Yoon, & S. S. Lam "Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms", *Expert Systems with Applications*, Vol. 41, No. 4, pp. 1476-1482, 2014.

[22] B. C. Patel, & G. R. Sinha "An adaptive K-means clustering algorithm for breast image segmentation", *International Journal of Computer Applications*, Vol. 10, No. 4, pp. 35-38, 2010.

[23] A. Sathish, and J. M. Sundaram. "A comparative study on K-means and fuzzy C-means algorithm for breast cancer analysis." *International Journal of Computational Intelligence and Informatics*, Vol. 4, No. 1, pp: 54-58, 2014.

[24] R. Preetha, & G. R. Suresh "Performance analysis of fuzzy c means algorithm in automated detection of brain tumor", *World Congress on Computing and Communication Technologies*, pp. 30-33, 2014.

***