International Journal of Early Childhood Special Education (INT-JECSE)
ISSN: 1308-5581 Vol 14, Issue 03 2022

ENHANCING ACCURACY IN DETECTION AND CLASSIFICATION OF WHITE BLOOD CANCER CELLS USING K-MEANS OVER MORPHOLOGICAL SEGMENTATION

BurlaGopi raju[1], S. Narendran[2]

[1]Research Scholar, Department of Electronics and Communication Engineering, Saveetha School of Engineering,Saveetha Institute of Medical And Technical Sciences, SaveethaUniversity, Chennai, Tamilnadu.India. Pincode: 602105.

[2]Project Guide, Corresponding Author, Department of Electronics and Communication Engineering, Saveetha School of Engineering, Saveetha Institute of Medical And Technical Sciences, Saveetha University, Chennai, Tamilnadu. India. Pincode: 602105.

E.mail :[1]burlagopiraju17@saveetha.com, [2]narendrans.sse@saveetha.com

ABSTRACT :Aim: The main objective of the work is to improve the accuracy and specificity for automatic detection, classification, and counting of the number of cells using a Morphological segmentation algorithm compared with the K-Means clustering algorithm. **Materials and Methods:** Data set used for this research consists of 28 cancer images for the detection of different types of White blood cell (WBC) cancer diseases. Determination of sample size of the two groups is calculated using G power (power of 0.80 and alpha value of 0.05) and two groups are categorized as Morphological segmentation classifier (Group 1) and k-means clustering classifier (Group 2). For performance analysis, 70% of the images are used for training and 30% of images are used for testing and validation. **Result:** Morphological segmentation algorithm achieved improved accuracy (97.77%) compared with the K-means clustering algorithm with an accuracy of (82.22%). The independent sample T-test has been analyzed and achieved a significance of 0.227 ($p > 0.05$) and specificity of 0.006 ($p < 0.05$). **Conclusion:** Morphological segmentation algorithm provides better accuracy compared with the K - means clustering algorithm.

KEYWORDS: Innovative White Blood Cancer Detection, Segmentation, Morphological Segmentation Algorithm, K-Means Clustering, Cancer Disease, White Blood Cells.

1. INTRODUCTION

Cancer is considered to be one of the most fatal diseases which can cause death. Around one in six deaths in the world are due to cancer disease (Agrawal et al. 2019). Automated classification and morphological analysis of white blood cells have been addressed for the last four decades, but there is no optimal method that can be used as a decision support system in laboratories due to the biologically complex nature of the cells (Hegde et al. 2019). At present, software simulations are widely used in medical imaging which includes segmentation, identification, and classification of cell/ nuclei. The microscopic blood image analysis finds its usefulness in differential and complete count of White blood cells (WBCs) and components, hemiparasite and hematological diseases diagnosis, content-based image retrieval systems, medical decision support systems, computer-aided diagnosis, treatment follow up guided surgery, and many more (Hegde et al. 2019; Tali, Borra, and Mahmud 2021).

In Google Scholar, 16,500 publications are published on segmentation and detection algorithms from (2016-2021). Some recent works related to WBCs and RBCs segmentation are based on color space such as HSI color space (Ismail and Moetesum 2018). The other methods used for conventional machine learning and image processing techniques are K-Means clustering, Fuzzy C-Means clustering, and Watershed Segmentation. Since the blood cells have a huge diversity in shape and color, the segmentation of exactly red blood cells and white blood cells in blood images is still a challenge (Ismail and Moetesum 2018; Kim et al. 2020). Image processing is concerned with digital images to extract useful information. It is involved in different other topics such as layout land use Character recognition coin recognition Medical imaging. Medical imaging uses the information extracted from digital images to enhance the diagnosis of different diseases. White blood cell cancer diseases; Leukemia and Myeloma, threaten people's lives nowadays. Leukemia is found when the bone marrow produces

DOI: 10.9756/INT-JECSE/V14I3.742

abnormal white blood cells, which don't function properly. In this article (Mohamed et al. 2018), the authors proposed an ALL detection system using a novel approach. They have used the marker-based segmentation (MBS) method for segmentation and GLCM for feature extraction. For feature reduction, they have used Probabilistic principal component analysis (PPCA) (Mishra et al. 2017). A texture-based feature extraction method for detection of cancer cells in joints. This method can be used in the early stages and it yields higher accuracy. The author used median filter and edge sharpening function for pre-processing of image (Arya, Mittal, and Singh 2018).Our team has extensive knowledge and research experience that has translate into high quality publications(Patturaja and Pradeep 2016; Ramesh Kumar et al. 2011; Krishnan, Pandian, and Kumar S 2015; Felicita 2017b, [a] 2017; Kumar 2017; Sivamurthy and Sundari 2016; Sathivel et al. 2008; Sekar et al. 2019)
In the existing method, two types of cancer cells were only detected. Counting cancer cells cannot be done. Here further development of the process is also needed to process the blood cancer. Proposed method provides a robust mechanism for the classification of Acute Lymphoblastic Leukemia (ALL) and Multiple Myeloma (MM) using the SN-AM dataset.

2. MATERIALS AND METHODS:

This study was carried out in the Image processing laboratory, Department of Electronics and Communication Engineering at Saveetha School of Engineering. The data set used for this experiment is obtained from the Kaggle open-access dataset. Sample size of groups are determined using G power (power value of 0.80 and alpha value of 0.05 (error rate) with inputs of 2400 (1200*2)). It consists of two groups and two sets of data have been collected. In this framework, 1200 images are used out of which (80%) are trained and (20%) are used for testing for the Innovative white blood cancer detection(Kg, Lakshmi, and ManjaNaik 2019). The proposed work is based on a computerized program in python using image samples of two groups. Total sample size per group is 28.
In Group 1, the cancer images are tested and performance analysis is done using a Morphological segmentation algorithm and similarly, in Group 2, the cancer images are tested and performance analysis is done using the k-means clustering algorithm in the Matlab platform.
This research work flows as follows as image preprocessing, Image segmentation, Classification, and Detection. According to this, the enteredimage is given to the machine and the Pre-processing consists of the image resizing and the development of the image data. Image segmentation is finishedby means ofthe usage ofautomatedthresholding segmentation and overlaying operations in red, green, and blue (RGB) planes. Threshold segmentation converts an imageright into a binary image (White or Black). Entropy, shape, strengthcapabilities are extracted. 70% of the pics from the dataset are used for schooling and 30% of images are used for trying out the modern detection of pores and skin cancer. It is a non-linear operation associated with the form and length or morphology functions withinside the photo. It successfully offers geometrical functions which include contrast, form, and length that may be taken into consideration as segmentation-orientated functions. Total of 48 white blood cell images are used for the classification of cancer images. Out of them, 28 are labeled "cancer white blood cells", which indicates that it is a cancer image and the remaining 20 are labeled as Healthy "white blood cells", which indicates that it is a healthy white blood cell. This proposed work involves cancer detection and classification for each group. The testing setup uses core i5 @ 1.65 GHZ, 4 GB RAM, X-64 based processor. MATLAB software with the installed library.

Statistical Analysis:

Independent t - test analysis is carried out using SPSS software. The input image intensity is an independent variable and the epoch values are dependent variables to determine the accuracy.

3. RESULTS:

Proposed system uses the Morphological segmentation algorithm for the detection and classification of white blood cell cancer. Morphological segmentation defines the Size and shape of the image and that combines morphological operations such as extended minima and morphological gradient and gives an accuracy of 97.77%, and specificity of 97.36%. Morphological segmentation shows better results compared with the k-means clustering algorithm. All the input images are resized into the same dimensions which is shown in Fig. 1. If the specified size does not produce the same aspect ratio as the input image, the output image will be distorted. Figure 2 shows clustering is an unsupervised machine learning method of identifying and grouping similar data points In the above three clusters, cluster 3 is having a better accurate vision. Figure 3 shows the mask is created to find the exact operations in an image. So that we can identify the problems or the features which we need to find in an image. The concept of masking is also known as spatial filtering.

Table 1 shows thesample dataset of SPSS software. Group 1 samples are obtained from morphological segmentation and Group 2 samples are obtained from k-means clustering. Comparison table shows a more accurate value for morphological segmentation and k-means clustering. Morphological segmentation algorithm achieved improved accuracy (97.77%) compared with the K-means clustering algorithm with an accuracy of (82.22%). Table 2 shows that the comparison of the two groups is made with parameters mean, Standard deviation, and standard error mean. The mean value is 97.4120 for morphological and 82.1230 for K - mean clustering algorithm for accuracy comparison. It declares that the Morphological segmentation performed better when compared with the k-means clustering algorithm. Table 3 shows Morphological segmentation algorithm achieved improved accuracy (97.77%) compared with the K-means clustering algorithm with an accuracy of (82.22%). The independent sample T-test has been analyzed and achieved a significance of 0.227 ($p > 0.05$) and specificity of 0.006 ($p < 0.05$) has been achieved.

In Fig. 4, the green color region indicates that white blood cancer is detected. At the time of disease detection, the total number of cells in the image will also be detected. In Fig. 4, the total number of cells detected in the CLL disease is 19.

In Fig. 5, the bar graph is plotted which represents the relationship between the two groups' mean accuracy with the error bar. In this graph, the value of Error Bars is 95% CI and +/-1 SD. The value of accuracy is 97.77% and specificity of 97.36% in group 1 and the value of accuracy is 82.22% and specificity of 81.57% in group 2. Based on this graph it concludes that morphological segmentation has more accuracy than the k-means clustering algorithm.

4. DISCUSSIONS:

In this framework, it is observed that the values of the mean accuracy using the morphological segmentation algorithm are high compared to the k-means clustering algorithm with a significance p-value greater than 0.05 in innovative white blood cancer detection. The statistical analysis produces the mean accuracy, standard deviation and standard error mean of morphological segmentation of accuracy is 97.4120, 0.25231 and 0.07979 and specificity is 97.2140, 0.12843 and 0.04061. Similarly the accuracy of K - means clustering is 82.1230, 0.07761 and 0.02454 and the corresponding specificity is 81.3010, 0.18570 and 0.05872. K-means clustering produces the least discrepancy of mean accuracy with significance value $p > 0.05$.

In the previous studies (Saritha et al. 2016), the authors demonstrated white blood cancer disease-based model architecture to classify cancer cell images. The performance is showing the least difference between the groups with a significance of less than 0.05. In this article the first step, white blood cells are identified using color space conversion models. Then leukocyte groups are separated using the division of watershed conversion. In the next step, image cleanup is done and all leukocytes available on the edge of images and abnormal components are removed. This is accomplished by cutting the image with the smallest rectangle that has connected components. The second level of division relates to the detection of the nucleus and cytoplasm. In the last step, feature extraction is performed which causes the pathologists to have the best interpretation of them and showed an accuracy of 93% in the detection of white blood cells (Alreza and Karimian 2016). In this article (Shrimali, Pandey, and Chowdhary 2021), the analysis is carried out using multiple machine learning algorithms. It was concluded that using neural networks is the best possible method to detect pulsar emissions from neutron stars. The best result achieved was 98% using Neural Networks. In this article (Shrimali, Pandey, and Chowdhary 2021; Loey, Naman, and Zayed 2020), the authors proposed SVM (support vector machine) algorithm and CNN (Convolutional Neural Networks) algorithm with the result of accuracy value it can be concluded that the proposed system CNN was achieving huge successes in the field whether regarding features extraction or classification task, time, accuracy and had a lower cost in the detection of leukemia diseases. In this study (Shrimali, Pandey, and Chowdhary 2021; Loey, Naman, and Zayed 2020; Su, Liu, and Song 2017), the authors proposed a method to segment acute leukemia and normal cells and their nuclei and cytoplasm from blood smear images and achieved an overall accuracy of 96% and 93% to discriminate between acute leukemia and normal cell region and their nucleus region, respectively. In this article (Yao et al. 2013), the authors proposed a new method to improve the stability and accuracy and have used mathematical morphology and segmentation by clustering using K-means. Also, they have converted the original image to grayscale and then used the improved K-means clustering algorithm to get better results.

Although the proposed method has better performance than the traditional method, it also has few limitations. Proposed technique offers pathologically validated facts and it overcomes the magnificence imbalance and overtraining issues, the overall performance degree may be progressed to a more quantity if they enter facts that have extra attributes. Overall performance metrics may be progressed via means of editing the present set of rules via means of growing or lowering the convolution stages. Proposed set of rules may be progressed via ways of thinking about entering attributes like age, gender, degree of observation, etc.

5. CONCLUSION:

The innovative detection of white blood cancer cell analysis of morphological segmentation shows better results compared with the k-means clustering algorithm. The deviation between those two groups are acceptable ranges hence, it is concluded that morphological segmentation algorithm and k-means clustering algorithm can be applied in the medical field for the early prediction of white blood cancer. An innovative detection of white blood cancer using a deep learning algorithm shows good performance in the medical field.

DECLARATIONS:

Conflict of Interest:

No Conflict of Interest in this manuscript.

Author contribution:

BG carried out the detection of white blood cancer studies, collected data, performance and analysis, statistical analysis, and drafted the manuscript. SN was involved in the study, coordinated, guided, and helped to draft the manuscript.

Acknowledgment:

The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical And Technical Sciences (Formerly known as Saveetha University) for providing the necessary infrastructure to carry out this work successfully.

Funding: We thank the following organizations for providing financial support that enabled us to complete the study.

1. VK Technologies, Trichy
2. Saveetha University
3. Saveetha Institute of Medical And Technical Sciences
4. Saveetha School of Engineering

6. REFERENCES:

[1] Agrawal, Rohit, Sachinandan Satapathy, Govind Bagla, and K. Rajakumar. 2019. “Detection of White Blood Cell Cancer Using Image Processing.” *2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN)*. https://doi.org/10.1109/vitecon.2019.8899602.

[2] Alreza, Zahra Khandan Khadem, and Alireza Karimian. 2016. “Design a New Algorithm to Count White Blood Cells for Classification Leukemic Blood Image Using Machine Vision System.” *2016 6th International Conference on Computer and Knowledge Engineering (ICCKE)*. https://doi.org/10.1109/iccke.2016.7802148.

[3] Arya, Mithlesh, Namita Mittal, and Girdhari Singh. 2018. “Texture-based Feature Extraction of Smear Images for the Detection of Cervical Cancer.” *IET Computer Vision*. https://doi.org/10.1049/iet-cvi.2018.5349.

[4] Felicita, A. Sumathi. 2017a. “Quantification of Intrusive/retraction Force and Moment Generated during En-Masse Retraction of Maxillary Anterior Teeth Using Mini-Implants: A Conceptual Approach.” *Dental Press Journal of Orthodontics* 22 (5): 47–55.

[5] ———. 2017b. “Orthodontic Management of a Dilacerated Central Incisor and Partially Impacted Canine with Unilateral Extraction - A Case Report.” *The Saudi Dental Journal* 29 (4): 185–93.

[6] Hegde, Roopa B., Keerthana Prasad, Harishchandra Hebbar, and Brij Mohan Kumar Singh. 2019. “Comparison of Traditional Image Processing and Deep Learning Approaches for Classification of White Blood Cells in Peripheral Blood Smear Images.” *Biocybernetics and Biomedical Engineering*. https://doi.org/10.1016/j.bbe.2019.01.005.

[7] Ismail, Basit, and Momina Moetesum. 2018. “Automated Detection and Quantification of Erythrocytes and Leukocytes from Giesma Stains of Blood Smear.” *2018 14th International Conference on Emerging Technologies (ICET)*. https://doi.org/10.1109/icet.2018.8603620.

[8] Kg, Lakshmi, K. G. Lakshmi, and N. Manja Naik. 2019. “Automated Detection of White Blood Cells Cancer Disease.” *2019 1st International Conference on Advances in Information Technology (ICAIT)*.

https://doi.org/10.1109/icait47043.2019.8987352.

[9] Kim, Byeonghwi, Yuli-Sun Hariyani, Young-Ho Cho, and Cheolsoo Park. 2020. “Automated White Blood Cell Counting in Nailfold Capillary Using Deep Learning Segmentation and Video Stabilization.” *Sensors* 20 (24). https://doi.org/10.3390/s20247101.

[10] Krishnan, Sindhuja, SaravanaPandian, and Aravind Kumar S. 2015. “Effect of Bisphosphonates on Orthodontic Tooth Movement-an Update.” *Journal of Clinical and Diagnostic Research: JCDR* 9 (4): ZE01–05.

[11] Kumar, Santhosh. 2017. “The Emerging Role of Botulinum Toxin in the Treatment of Orofacial Disorders: Literature Update.” *Asian Journal of Pharmaceutical and Clinical Research* 10 (9): 21.

[12] Loey, Mohamed, Mukdad Rasheed Naman, and Hala Helmy Zayed. 2020. “A Survey on Blood Image Diseases Detection Using Deep Learning.” *International Journal of Service Science, Management, Engineering, and Technology*. https://doi.org/10.4018/ijssmet.2020070102.

[13] Mishra, Sonali, Banshidhar Majhi, Pankaj Kumar Sa, and Lokesh Sharma. 2017. “Gray Level Co-Occurrence Matrix and Random Forest Based Acute Lymphoblastic Leukemia Detection.” *Biomedical Signal Processing and Control*. https://doi.org/10.1016/j.bspc.2016.11.021.

[14] Mohamed, Hend, Rowan Omar, Nermeen Saeed, Ali Essam, Nada Ayman, Taraggy Mohiy, and Ashraf AbdelRaouf. 2018. “Automated Detection of White Blood Cells Cancer Diseases.” *2018 First International Workshop on Deep and Representation Learning (IWDRL)*. https://doi.org/10.1109/iwdrl.2018.8358214.

[15] Patturaja, Kiruthika, and D. Pradeep. 2016. “Awareness of Basic Dental Procedure among General Population.” *Journal of Advanced Pharmaceutical Technology & Research* 9 (9): 1349.

[16] Ramesh Kumar, K. R., K. K. ShantaSundari, A. Venkatesan, and ShymalaaChandrasekar. 2011. “Depth of Resin Penetration into Enamel with 3 Types of Enamel Conditioning Methods: A Confocal Microscopic Study.” *American Journal of Orthodontics and Dentofacial Orthopedics: Official Publication of the American Association of Orthodontists, Its Constituent Societies, and the American Board of Orthodontics* 140 (4): 479–85.

[17] Saritha, M., B. B. Prakash, K. Sukesh, and B. Shrinivas. 2016. “Detection of Blood Cancer in Microscopic Images of Human Blood Samples: A Review.” *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*. https://doi.org/10.1109/iceeot.2016.7754751.

[18] Sathivel, Arumugam, Hanumantha Rao BalajiRaghavendran, Periasamy Srinivasan, and ThiruvengadamDevaki. 2008. “Anti-Peroxidative and Anti-Hyperlipidemic Nature of UlvaLactuca Crude Polysaccharide on D-Galactosamine Induced Hepatitis in Rats.” *Food and Chemical Toxicology: An International Journal Published for the British Industrial Biological Research Association* 46 (10): 3262–67.

[19] Sekar, Durairaj, Ganesh Lakshmanan, Panagal Mani, and M. Biruntha. 2019. “Methylation-Dependent Circulating microRNA 510 in Preeclampsia Patients.” *Hypertension Research: Official Journal of the Japanese Society of Hypertension* 42 (10): 1647–48.

[20] Shrimali, Shubham, Amritanshu Pandey, and Chiranji L. Chowdhary. 2021. “K-Means Clustering-Based Radio Neutron Star Pulsar Emission Mechanism.” *Recent Advances in Computer Science and Communications*. https://doi.org/10.2174/2213275912666200129115401.

[21] Sivamurthy, Gautham, and ShanthaSundari. 2016. “Stress Distribution Patterns at Mini-Implant Site during Retraction and Intrusion—a Three-Dimensional Finite Element Study.” *Progress in Orthodontics* 17 (1): 1–11.

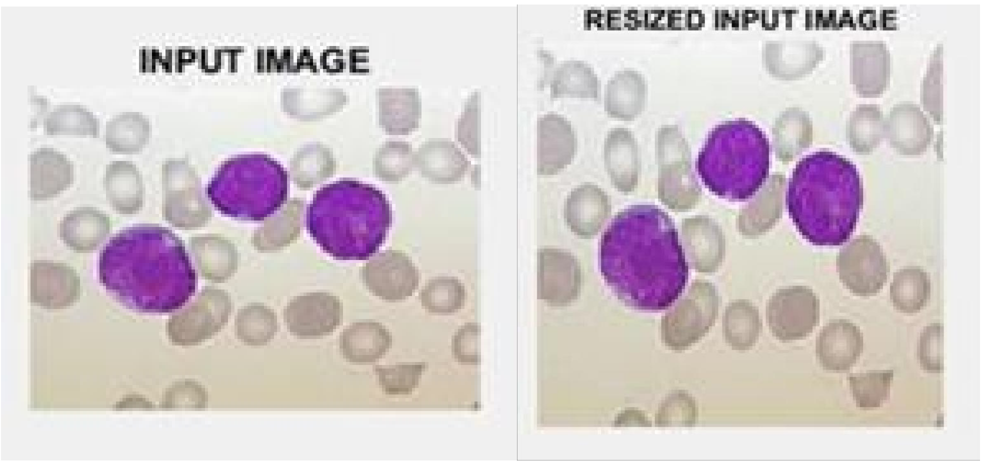
[22] Su, Jie, Shuai Liu, and Jinming Song. 2017. “A Segmentation Method Based on HMRF for the Aided Diagnosis of Acute Myeloid Leukemia.” *Computer Methods and Programs in Biomedicine* 152 (December): 115–23.

[23] Tali, Renuka Veerappa, Surekha Borra, and Mufti Mahmud. 2021. “Detection and Classification of Leukocytes in Blood Smear Images.” *International Journal of Ambient Computing and Intelligence*. https://doi.org/10.4018/ijaci.2021040107.

[24] Yao, Hong, Qingling Duan, Daoliang Li, and Jianping Wang. 2013. “An improvedK-Means Clustering Algorithm for Fish Image Segmentation.” *Mathematical and Computer Modelling*. https://doi.org/10.1016/j.mcm.2012.12.025.

[25]

FIGURES AND TABLES



(a) **(b)**

Fig. **1.** Input images for segmentation and classification: a) Input Image, b) Resized Image

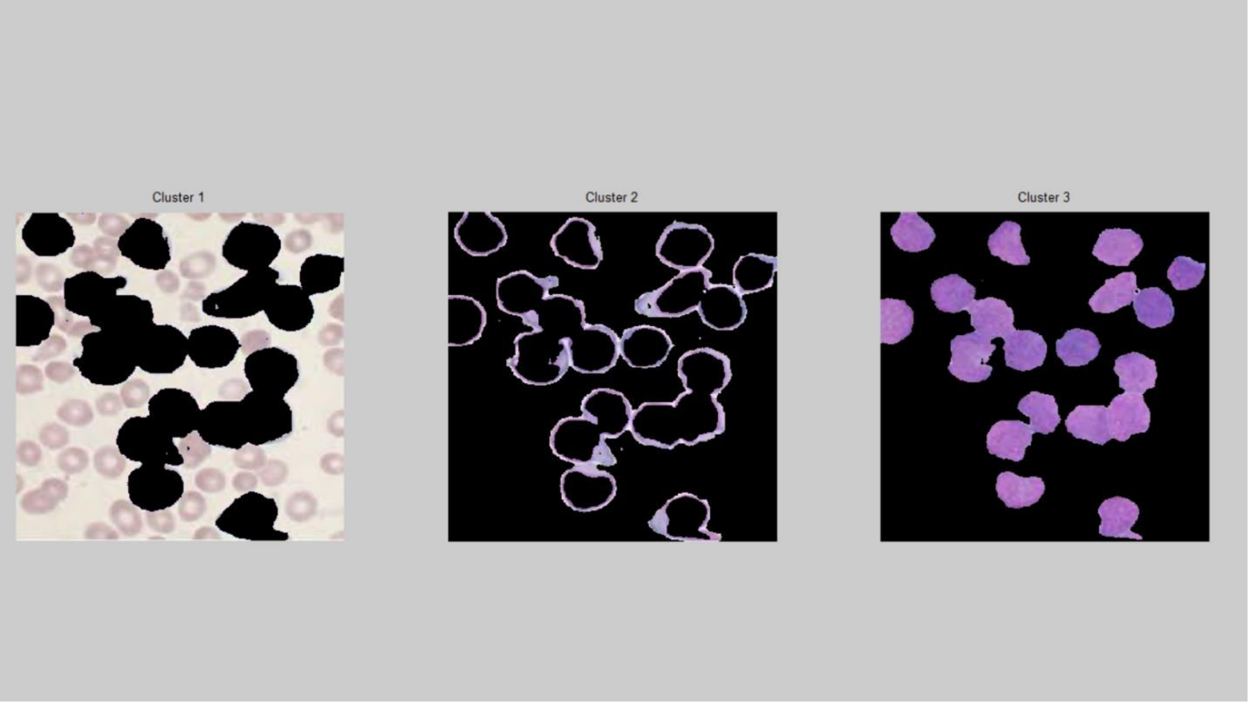


Fig. 2. Clustering Image by grouping similar data.

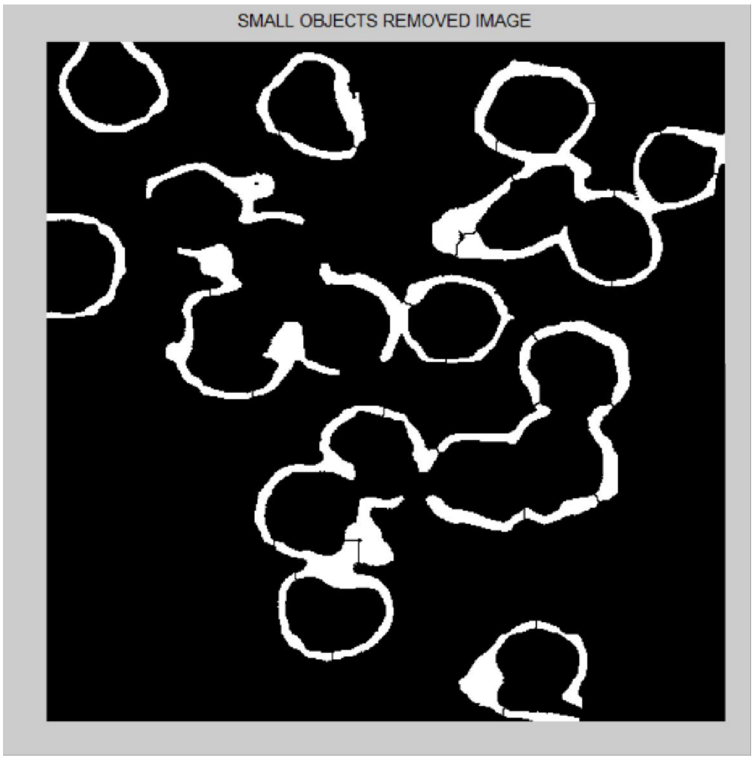


Fig. 3. Masking the Image after clustering

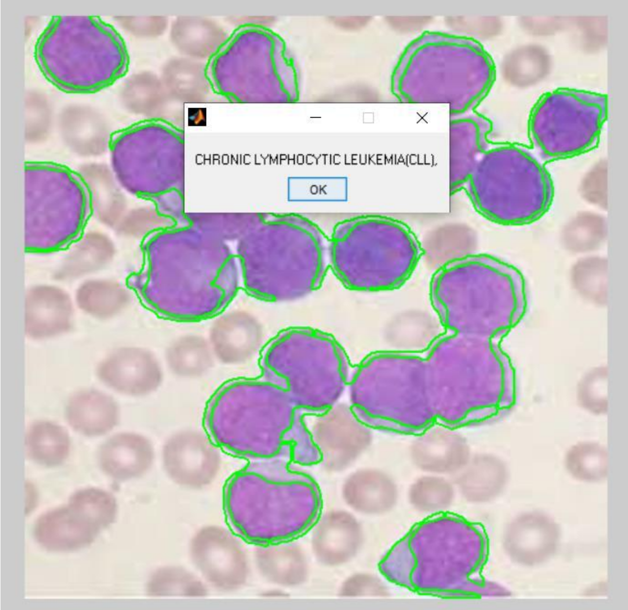

CHRONIC LYMPHOCYTIC LEUKEMIA(CLL).

Fig. 4. The total number of cells detected in the CLL disease is 19.

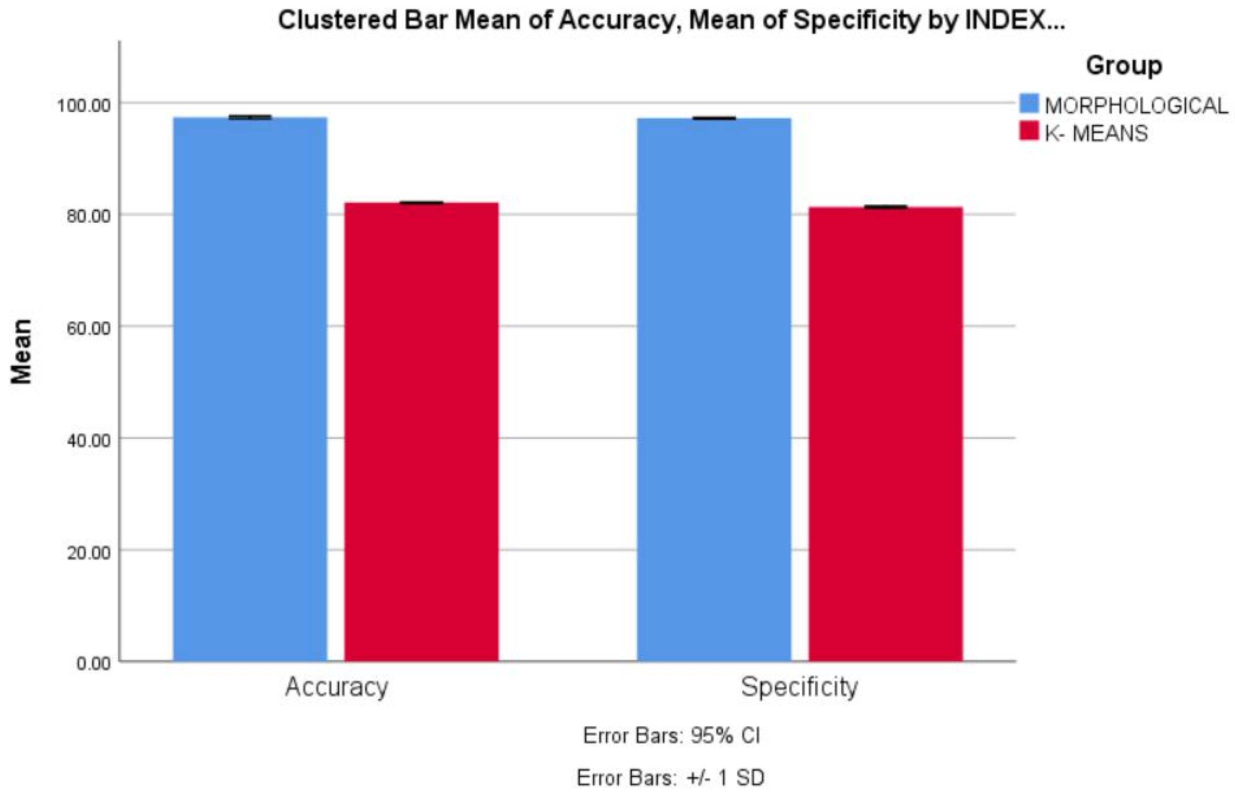


Fig. 5. Bar graph analysis of two groups (Morphological and K - means clustering methods) based on accuracy and specificity. The X-axis represents Group 1 and 2 Algorithms: Y-axis represents Mean Accuracy, specificity with +/- 1 SD.

Table 1. Sample Dataset of SPSS Software. Group 1 samples are obtained from morphological segmentation and Group 2 samples are obtained from k-means clustering. The Comparison shows a more accurate value for morphological segmentation than k-means clustering.

S.No	Group	Sample	Accuracy	Specificity
1	1	Sample 1	97.77	97.36
2	1	Sample 1	97.67	97.32
3	1	Sample 1	97.50	97.30
4	1	Sample 1	97.45	97.27
5	1	Sample 1	97.40	97.24
6	1	Sample 1	97.30	97.20
7	1	Sample 1	97.32	97.17
8	1	Sample 1	97.22	97.14
9	1	Sample 1	97.15	97.08

10	1	Sample 1	97.00	97.00
11	2	Sample 2	82.22	81.57
12	2	Sample 2	82.20	81.50
13	2	Sample 2	82.17	81.45
14	2	Sample 2	82.15	81.40
15	2	Sample 2	82.13	81.36
16	2	Sample 2	82.10	81.30
17	2	Sample 2	82.07	81.25
18	2	Sample 2	82.05	81.17
19	2	Sample 2	82.03	81.11
20	2	Sample 2	82.00	81.03

Table 2. SPSS Statistical Analysis of k-means clustering and morphological segmentation models. Mean Accuracy, Specificity, Standard Deviation and Standard Errors Mean are obtained for 10 iterations. The performance represents better outcomes for morphological segmentation than the k-means clustering.

Group Statistics					
	Group	N	Mean	Std.Deviation	Std.Error Mean
Accuracy	Morphological segmentation	10	97.4120	0.25231	0.07979
	K-Means clustering	10	82.1230	0.07761	0.02454
Specificity	Morphological segmentation	10	97.2140	0.12843	0.04061
	K-Means clustering	10	81.3020	0.18570	0.05872

Table 3. Independent sample T - test is performed for the two groups for significance and standard error determination.Levene's Test for Equality of variances showed a significance of 0.227 ($p > 0.05$) for accuracy and 0.006 ($p < 0.05$) for specificity.

Levene's Test for Equality of Variances				T-test for Equality of Means					95% Confidence Interval of the Difference	
F			Sig	t	df	Sig(2-tailed)	Mean Difference	Std.Error Difference	Lower	Upper
Accuray	Equal variances assumed	1.568	0.227	222.861	18	0	15.912	0.0714	15.762	16.062
	Equal variance not assumed			222.861	16.006	0	15.912	0.0714	15.7607	16.0634
Specificity	Equal variances assumed	9.691	0.006	183.15	18	0	15.289	0.08348	15.1136	15.4644
	Equal variance not assumed			183.15	10.688	0	15.289	0.08348	15.1046	15.4734

DOI: 10.9756/INT-JECSE/V14I3.742