ORIGINAL ARTICLE

INTERNATIONAL JOURNAL OF RESEARCH IN PHARMACEUTICAL SCIENCES

Published by JK Welfare & Pharmascope Foundation Journal Home Page: https://ijrps.com

An efficient technique for identification of leukemia in microscopic blood samples using image processing

Meghana M.R*, Akshatha Prabhu

Department of Computer Science, Amrita School of Arts and Sciences, Mysuru Amrita Vishwa Vidyapeetham, karnataka, India

Article History:

Received on: 17.04.2019
Revised on: 05.07.2019
Accepted on: 09.07.2019

Keywords:

leukemia,
white blood cells,
bone marrow,
blood cancer,
classified

ABSTRACT

Leukemia is a blood cancer which features through the ejection of manipulated and strange fabrication of white blood cells which is the way of bone marrow within the blood. The project aims at designing and developing an efficient technique for the detection of luekemia based on image segmentation techniques and nuclei analysis which incorporates the affected percentage and are compared and classified using KNN and SVM. The DNA of youngster cells, for the maximum detail white platelets, subsequently finally ends up harmed here and there. This version from the norm reasons platelets to increase and separate constantly. Sound platelets bypass on inevitably and are supplanted by approach of new cells, which might be brought in bone marrow.

*Corresponding Author

Name: Meghana M.R
Phone: 9611402747
Email: meghanagowda077@gmail.com

ISSN: 0975-7538

DOI: https://doi.org/10.26452/ijrps.v10i3.1487

Production and Hosted by

IJRPS | https://ijrps.com

INTRODUCTION

Leukemia is a malignant blood increase. The word Leukemia originates from the Greek 'leukos', which indicates 'white' and 'aima', which indicates 'blood'. The DNA of youngster cells, for the maximum detail white platelets, subsequently finally ends up harmed here and there. This version from the norm reasons platelets to increase and separate constantly. Sound platelets bypass on inevitably and are supplanted by the approach of new cells, which might be brought in the bone marrow. The bizarre cells do not kick the bucket when they want to and acquire, regarding greater space. As an increasing number of malignant increase cells are delivered, they ward off the potential and development of sound white platelets thru jamming out the region in the blood. In the event that the visible example evaluation is automatic, it will help the pathologists with growing profitability and decrease charges. The robotization technique consists of image acquiring, image dealing with, department, spotlight extraction and association. The division is taken into consideration the maximum critical and simple develop in the technique because it affects the rest. Blood cell branch basically eliminates the telephones from stressed foundation and fragments each telephone into morphological segments, for example, centre cytoplasm, and a few others. Here errand is to perceive extra younger mobile making use of one-of-a-type picture getting ready strategies and tally the whole variety of cells. So we've had been given to utilize the innovation that distinguishes various. Forms of platelets interior quick span of time in the crisis. Moreover, it's far critical to contemplate in detail a way to cut up diverse mobile and recollect it as younger cell and as indicated through it, recognize leukemia. In leukemia is notion to arise whilst a few blood cells acquire mutations in their DNA the commands inside every mobile that guide its motion. Certain abnormalities reason the cellular to grow and divide more swiftly and to hold living whilst regular cells could die. Over time

these odd cells can crowd out wholesome blood cells within the bone marrow main to fewer healthy WBC, RBC and platelet inflicting the symptoms and signs of leukemia. No one knows exactly what causes leukemia. People who have it has certain atypical chromosomes; however, the chromosomes don't motive leukemia.

Literature survey

(Ongun *et al.*, 2001) this paper present a comprehensive survey approximately segmentation methods. The primary goal of his examine is to develop an automatic device on blood cell class. His paintings summarize the most famous and normal methodologies useful for the assessment of image analysis prominently in the segmentation technique. It segments the WBCs the use of active contour models.

(Shankar *et al.*, 2016) Detects the extreme lymphoblastic leukemia that's the form of most cancers due to the production of immature cells this is in the main affected children underneath five years and adults above 50 years age. Detection of it is recognized and considers the infected WBC observed in human blood pattern. The method provides excessive speed, the accuracy of 90% and the scope for early detection.

(Khobragade *et al.*, 2015) Detects leukemia in microscopic bloodmobile photograph. The proposed approach is accomplished to the 90 microscopic blood cellular images. The algorithm is viable with Lab VIEW and IMAQ toolbox. It offers outcomes and accuracy of 85%.

(Rejintal and Aswini, 2016) the primary purpose of the paper is cellular segmentation which is continued by feature extraction, which locates forming cells. To consider for correct precisions of identification angular second moment, comparison, dissimilarity, homogeneity cluster prominence and inverse difference moment etc. are taken as the features. The outcomes display the K-means approach is applied for first-rate segmentation functionality.

(Tyagi *et al.*, 2016) the technique proposed on this paper is applicable nearly 100 images, and it exhibits uniformed result for all the image. It is been located that the segmentation of the blood corpuscles performs a vital feature in figuring out the kind of poikilocyte cells. This method affords 89% of accuracy.

Ni *et al.* (2016) we investigated the capacity of SVM to analyze MRD in drift cytometry records from sufferers with AML robotically, objectively and standard. Automated analysis outcomes-based at the version did no longer vary from and had been correlated with effects received thru conventional evaluation. Thus the SVM version may want to probably be used to investigate float cytometry-based totally AML MRD data robotically objectively and in a standardized way.

(Vogado *et al.*, 2016) Segmentation is taken into consideration a critical step inside the automatic prognosis of different computes. The images provided in this paper describe the brand new leukemic cells. Based on the effects obtained through the technique, it's far feasible to examine its effectiveness in contrast compared to some other works. Accuracy completed is 82% with rapid overall performance and hence a good accuracy.

(Gumble and Rode, 2017) The technique proposed in the images pursuits to diagnose leukemia the use of blood smear images. It's far possible to study that the device uses an image with none pre-processing and segmentation as enter. The principle difference among the method and the state-of-the-art techniques. SVM is used in the class to classify the images as pathological or not.

(Macawile *et al.*, 2018) this study became capable of classify and depend on white blood cells primarily based on microscopic blood samples with the aid of making use of CNN. Upon that, the contrast of the 3 models. It was found that Alex net performed excellent at the project to classify and based on counting totally on 21 blood pattern images when compared to google net and resnet-101 producing overall accuracy of89.18% and specify of 97.85% and the major accuracy of 96.63%.

(Putzu *et al.*, 2014) In this paper, various strategies used for detection of blood cancer automatically are reviewed. As blood cancer is one of the most dangerous ailment caused because of exceeds inside the number of white blood cells. Due to lack of proper treatment, it could be a deadly sickness. So it needs to be prognosis early.

(Pilzer *et al.*, 2010) the paper is implemented the cells which are exhibited increases resistance to MAC. It is a vital expressed heat shock protein concentrating on the binding of mortal in, analyzing the proteins, collection of secreted vesicles. It achieves the accuracy of 85 and 92% with respect to the methods used.

PROPOSED METHODOLOGY

Image Acquisition and Pre-processing

Because the blood scells image's length is seen via microscopic visuals. The input of the tool which is proposed are stained microscopic blood smear. The datasets of the microscopic luekemia are numerous. Among those, the equal antique ALL-IDB benchmark

dataset is choose for the research in this proposed paper. Then, the Processing the gained unformed snapshots aren't always feasible because of the fact that it may be of any other length. So for the requirement of solidarity inside the size. The images are transformed into the duration 256 × 256 matrix. Thus, to keep away from the complexity of visualization and for the processing, the image which is 3d that is RGB image has been modified into the grayscale image. The specification of the smallest and biggest stretching the limits of intensity values all the constituent pixels of the image are saved between 0 and 255. Here the pixel depth stretching tool gives greater outstanding exposure to the blood cells from their past history. The contrast of the image is a crucial thing to beautify the scene first-class of image. Hence, the evaluation of the blood smear image is an advanced technique of histogram equalization, which is used to adjust the image intensities to enhance contrast and can also overcome the overlapping cells as shown in Figure 1.

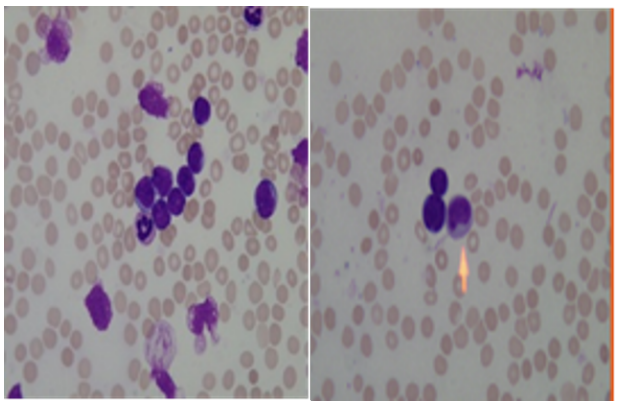

Figure 1: Original image

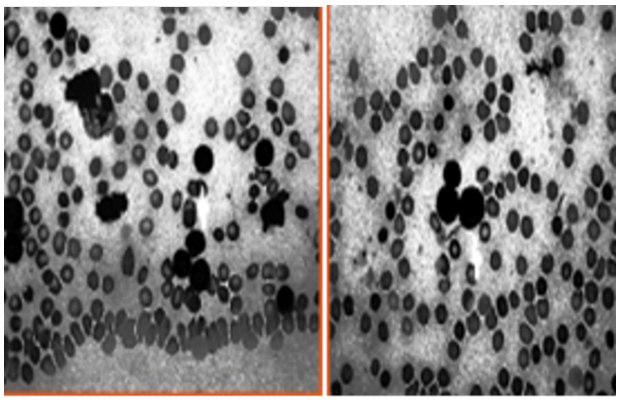

Figure 2: Histogram equalized

Histogram equalization

It is a technique for adjusting image intensities to enhance contrast. Let f be a given image represented as a matrix of integer pixel intensities ranging as shown in Figure 2.

Image Segmentation

Once the pre-processing step is completed, the intensity stretched and evaluation-better image are connected collectively because virtual images are segmented with the consideration of the matrices of the pixels, every detail of the primary image matrix is introduced to the further similar detail in the secondary image matrix. If the ensuring matrix maximizes the shape of the fractional numbers, and they will be rounded off. It brightens of all of the blood additives inside the image beside the nucleus of white blood cells. Every pixel of the image which is brighter matrix is subtracted using a manner of the corresponding pixel inside the assessment image, which is stretched to extract the bounds of blood additives. To extract the location of the ROI, the brightened image pixels are brought to correlate with the pixels within the extracted boundary image. Vicinity incorporates one valued pixel which might be black in coloration. The ultimate location filled up with zero-valued pixels which can be white in coloration. After that, to make nucleus and historical past in white and seem black in colorings, the picture has to be complimented. Correspondingly this will spotlight the place of interest, as shown in Figure 3.

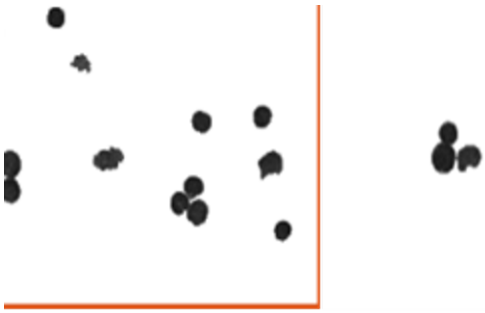

Figure 3: Final segmented image

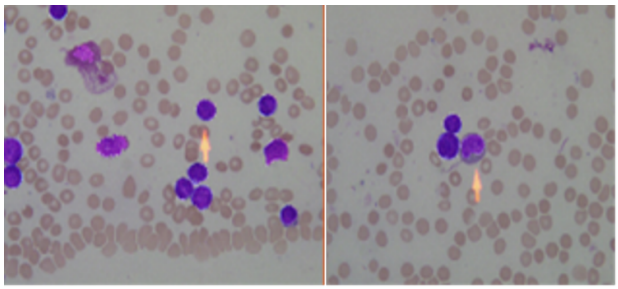

Figure 4: Separation of leukocytes

Separation of leukocytes

The leukocytes are separated from all other overlapped cells in order to find the effected cell, as shown in Figure 4.

Otsu's Thresholding

Thresholding helps to create a binary image that is white and black. So it is figured out that threshold fee which reduces weight inside elegance variance or increases the weight. This occurs operation immediately on greyscale images. In this method, we comprehensively search for a threshold, which reduces the intra-class variance that is the variance inside the class and named as a weighted sum of variances of the two different classes. It is usually a one-dimensional image. Using Otsu's methodology, we can calculate the optimum threshold, which separates the two classes, as shown in Figure 5. Hence, the motive of the Otsu's approach.

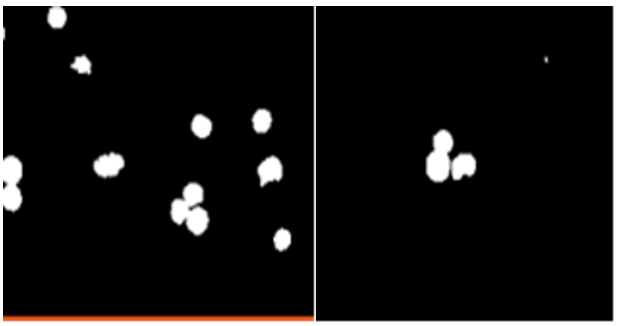

Figure 5: Segmentation using Otsu's method

Feature extraction and Classification

This is the phase where the features we want is extracted using a certain method in Matlab, and then it is classified in the respective classes or groups it belongs. The traits of a selected image item are located with the aid of extracting its competencies. Here, GLCM texture abilities, statistical and geometrical features of the nucleus are absorbed and exhibited to enter the method C-KNN classifier. C-KNN classifier measures the space among the capabilities of ROI, the usage of a custom-built distance metric, as demonstrated in an Equation (11). The blood image is classified into benign and malignant. The process flow of the method is proposed in Figure 6.

Figure 6: Affected by leukemia & healthy cell

Steps in proposed methodology

Phase 1

1. To get examined, obtain the microscopic blood smear image.
2. RGB image of 256×256 matrix is obtained.
3. The RGB image is converted into gray scale
4. Using linear contrast enhancement increase the depth of greyscale.
5. Using histogram equalization adjusts the contrast of the image.

$$K_{(Gray\ Image)} = 0.2988R + 0.5871G + 0.1142B \quad (1)$$

$$K_{(Intensity\ Enhanced)} = \frac{(K - Min) \times New\ Max - New\ Min + New\ Min}{Max - Min} \quad (2)$$

$$K_{(Contrast\_Stretched)}[L(i)] = ni/n,\ 0 \leq i < \quad (3)$$

Here ni is the entire number of occurrences of pixel i in an image and the total gray degrees inside the picture is I

Phase 2

6. Except the nucleus, brighten the image

$$K1 = K_{(Intensity\ Enhanced)} + K_{(Contrast\ Stretched)} \quad (4)$$

7. Borders of all the components are fragments.

$$K2 = K1 + K_{(Contrast\ Stretched)} \quad (5)$$

8. Uncouple the nucleus region of White blood cells.

$$K3 = K1 + K2 \quad (6)$$

Use 2-D order statistic minimum filtering, to filter an image

$$K_{(Filtered\ Image)}[L(i)] == Min\{IX1, \ldots, IXn\},\ 1 \leq i \leq n; \quad (7)$$

10. Binarize the filtered image using Otsu's global thresholding and complement to dark the required ROI.

Phase 3

11. Boundary of the selected region is dis-integrated and highlights invisible holes in the ROI

$$K_{(Eroded\ Image)} = K_{(Filtered\ Image)} \theta\ H = (\ \oplus\ H*)' \quad (8)$$

The structuring element is R

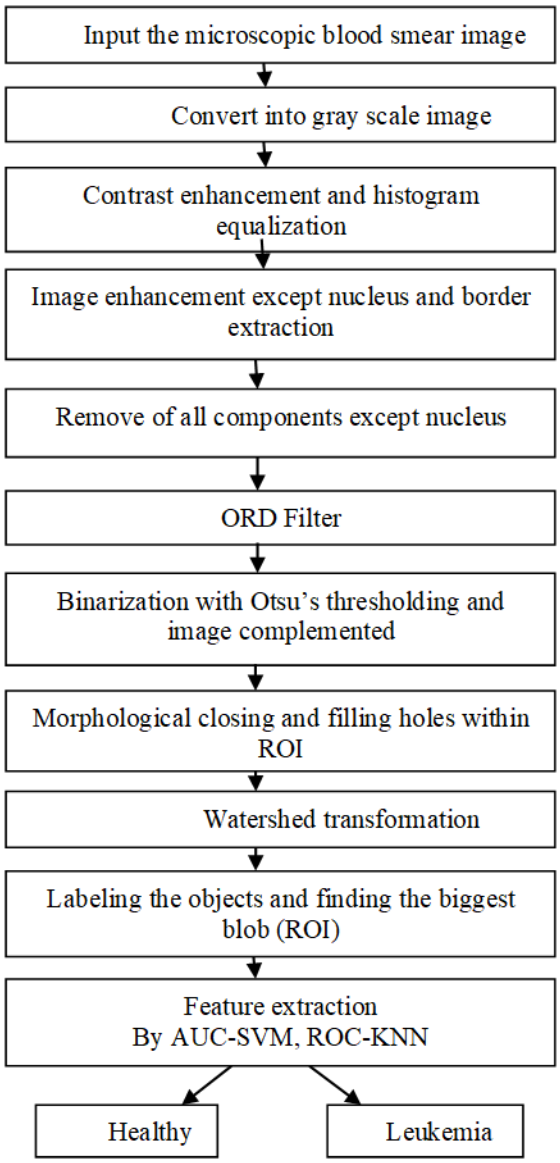

Chart 1: Flow chart of the proposed method

12. Close region which are selected by closing.

$$K_{(Closed\ Image)} = K_{(Eroded\ Image)} \cdot R = (K \oplus R) \ominus R \quad (9)$$

The structuring element is R

13. Fill the small holes inside the region

$$K_{(Holes\ Filled)}\ ROIK = (ROIK - 1 \oplus R) \cap K,\ for\ I = 1, 2, \ldots \quad (10)$$

14. Labialize the objects in an image to find the biggest ROI (blob)

Phase 4

15. The nucleus of features which are required are extracted.

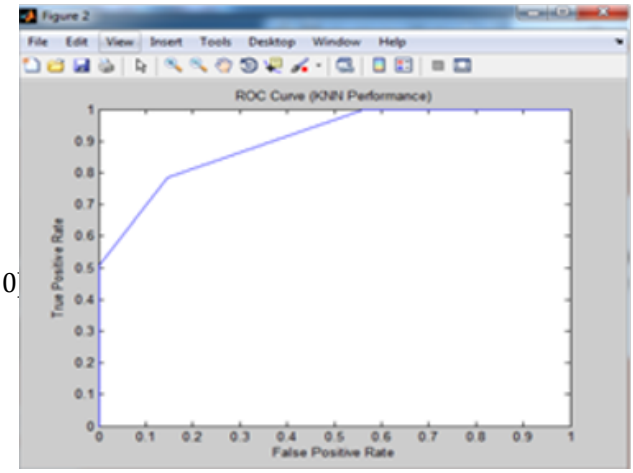

Figure 7: ROC curves using KNN classification

Table 1: The accuracy of proposed work compared to the existing techniques

Authors	Test images used	Classifiers	Accuracy
(Joshi *et al.*, 2013)	109	KNN	92%
(Rawat *et al.*, 2015)	265	SVM	89.9%
(Basima and Panicker, 2016)	92	SVM	93.56%
(Bhukya and Somayajulu, 2011)	85	SVM	94.38%
(Kumar *et al.*, 2011)	61	KNN	91.8%
(Moshavash *et al.*, 2018)	109	KNN	90.76%
Approach proposed with the K-NN (applied Euclidean) (2018)	82	KNN-Euclidean	94.96%
approach proposed with C_ KNN (2018)	80	C-KNN	94%
Proposed work (2019)	103	ROC_KNN	96.02%

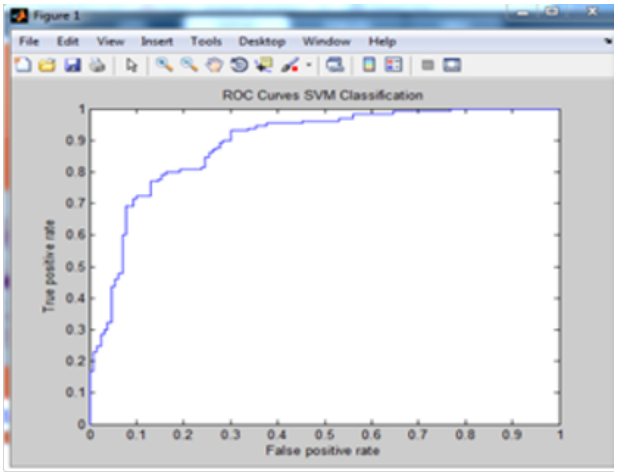

Figure 8: ROC curves using SVM classification

Figure 9: Performance Analysis using KNN and SVM

16. Images are classified using Customized KNN algorithm with the distance measure given below

$$S = (a + b) \tag{11}$$

Here, features are represented as a, b.

RESULTS AND DISCUSSION

The method which is proposed is experimented from ALL-IDB2 open source benchmark database measured 2592*1944 pixel resolution by using microscopic thin blobs smear images of 150 images. Using canon power shot G5 camera the images are captured which are attached to the microscopic optical laboratory. Inside the database, there are 130 malignant and 130 benign images of the entire category. The dataset is divided into a training dataset and test dataset. Over about 180 images are learned, and 80 images are tested by the system. The first 90 images from malignant and benign type to avoid the confusion in selecting the primary 90 images and Chosen for training and the rest of the 40 images with respect to the class or groups which are treated to test. The texture features are provided by the details about the difference between the pixels and its intensity as shown in Figure 7 and Figure 8. GLCM is the most important statistical texture extraction. By means of geometrical and texture features the nucleus region based on its visibility of the characteristics is analyzed. Here we will randomly choose. Each of their features are exhibited for illustration. Features are extracted from the same set of all 260 images, and later the datasets are trained and tested. Then they are selected to Classify cancerous and noncancerous images to the classifier C-KNN, as shown in figure g, h. The variety of gray levels of the image is reduced from to reduce the computational

time of GLCM in the matrix 32 to 256 levels through dividing the resized image with the correct value. Image = (256*256)/8(12). By identifying the resemblance and dissimilarity within the 103 test images features and compared with existing techniques in Table 1.

Here the perform analysis is done using K-NN and Support vector machine and the accuracy is achieved more using K-Nearest Neighbor technique as shown in Figure 9. The two different methods are used in order to compare the result which serves the purpose of the paper.

CONCLUSION

Our technique distinguishes the Leukemia with 77.25% exactness rate, which is superior to in different existing methodologies. Thus, it'll suggest step by step exact. Expectancies with fringe blood smear pictures. For doing this, the framework portions the centre of White Blood Cells that are ordered dependent on Gray Level Co-occurrence Matrix (GLCM) floor, measurable and geometrical highlights. As a completely novel method, the C-KNN classifier orders 97.25% precision affectability and 97.50% explicitness, as shown in figure e. This is 0.29% extra specific than K-Nearest Neighbor (KNN) and support vector machine (SVM) methods with Euclidian measure and a couple of 92% greater specific than the actual work. The AUC esteem done through this proposed system is 0.9725. In order to spare people lives, the propound technique can be utilized in hemato-pathology studies facilities, leukemia is figured out in order to oblige hematologists. Since this exploration proposition is delivering honorable outcomes, it'll cross about as a hobby for making changes inside the K-nearest neighbor calculation with other separation measures, as shown in figure i. So In the decided future work, the proceeding to perform upgrades utilizing the organization of diverse classifiers joined with the proposed framework to create drastically regularly specific outcomes.

FUTURE WORK

Future work which incorporates classification of lymphoblast into diverse sub types, counting and detecting different ordinary cells. In the future, someone can add better GUI design to look better and to implement this code or can use a few different algorithms to put in force the identical layout or calculate a few different parameters to enhance the accuracy and can identify the stage of the leukemia cancer a person is suffering. The paper can be further improved by classifying cancer affected person is male or female because the impact of cancer plays a different role in different genders though the symptoms are the same.

REFERENCES

Basima, C. T., Panicker, J. R. 2016. Enhanced leucocyte classification for leukaemia detection. *International Conference on Information Science (ICIS)*, pages 65–71.

Bhukya, R., Somayajulu, D. 2011. Exact Multiple Pattern Matching Algorithm using DNA Sequence and Pattern Pair. *International Journal of Computer Applications*, 17(8):32–38.

Gumble, P. M., Rode, S. V. 2017. Analysis & Classification of Acute Lymphoblastic Leukemia using KNN Algorithm. *International Journal on Recent and Innovation Trends in Computing and Communication*, 5(2):94–98.

Joshi, M. D., Karode, A. H., Suralkar, P. S. R. 2013. White Blood Cells Segmentation and Classification to Detect Acute Leukemia Ms. *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, 2(3):147–151.

Khobragade, S., Mor, D. D., Patil, C. Y. 2015. Detection of leukemia in microscopic white blood cell images. *International Conference on Information Processing (ICIP)*, pages 435–440.

Kumar, S., Panda, A. K., Singh, R. K. 2011. A review on tertiary recycling of high-density polyethylene to fuel. *Resources, Conservation and Recycling*, 55:893–910.

Macawile, M. J., Quinones, V. V., Ballado, A., Cruz, J. D., Caya, M. V. 2018. White blood cell classification and counting using convolutional neural network. *3rd International Conference on Control and Robotics Engineering (ICCRE)*, pages 259–263.

Moshavash, Z., Danyali, H., Helfroush, M. S. 2018. An Automatic and Robust Decision Support System for Accurate Acute Leukemia Diagnosis from Blood Microscopic Images. *Journal of Digital Imaging*, 31(5):702–717.

Ni, W., Hu, B., Zheng, C., Tong, Y., Wang, L., Li, Q., Han, Y. 2016. Automated analysis of acute myeloid leukemia minimal residual disease using a support vector machine. *Oncotarget*, (44):7–7.

Ongun, G., Halici, U., Leblebicioglu, K., Atalay, V., Beksac, M., Beksac, S. 2001. An automated differential blood count system. *2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 3:2583–2586.

Pilzer, D., Saar, M., Koya, K., Fishelson, Z. 2010. Mortalin inhibitors sensitize K562 leukemia cells

to complement-dependent cytotoxicity. *International Journal of Cancer*, 126(6):1428–1435. NA-NA.

Putzu, L., Caocci, G., Ruberto, C. D. 2014. Leucocyte classification for leukaemia detection using image processing techniques. *Artificial Intelligence in Medicine*, 62(3):179–191.

Rawat, J., Singh, A., Bhadauria, H. S., Virmani, J. 2015. Computer Aided Diagnostic System for Detection of Leukemia Using Microscopic Images. *Procedia Computer Science*, 70:748–756.

Rejintal, A., Aswini, N. 2016. Image processing based leukemia cancer cell detection. *IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTE-ICT)*, pages 471–474.

Shankar, V., Deshpande, M. M., Chaitra, N., Aditi, S. 2016. Automatic detection of acute lymphoblasitc leukemia using image processing. *IEEE International Conference on Advances in Computer Applications (ICACA)*, pages 186–189.

Tyagi, M., Saini, L. M., Dahyia, N. 2016. Detection of Poikilocyte cells in Iron Deficiency Anaemia using Artificial Neural Network. *International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC)*, pages 108–112.

Vogado, L. H. S., Veras, D. M. S., Andrade, R., Silva, A. R., Araujo, F. H. D., Medeiros, F. N. S. 2016. Unsupervised Leukemia Cells Segmentation Based on Multi-space Color Channels. *IEEE International Symposium on Multimedia (ISM)*, pages 451–456.