DETECTION OF WHITE BLOOD CELL CANCER DISEASES THROUGH CLASSIFICATION TECHNIQUES

N.D. Jambhekar*
Asst. Professor, Department of Computer Science, Gopikabai Sitaram Gawande Mahavidyalaya, Umarkhed, MS, India
P.S. Joshi
Asst. Professor, Department of Zoology, Shri Shivaji Arts, Commerce & Science College, Akot, MS, India
*jambhekar@gsgcollege.edu.in

Abstract
A biomedical research point is the computerised analysis of white platelet tumour diseases like Leukemia and Myeloma. A patient's health can be greatly improved by detecting problems early. In the event of a cancer discovery, blood smear images are a reliable and accurate source of information. It is imperative that feature extraction and reduction be done correctly when evaluating cancer growth finding such leukaemia cancer image. Diagnosing blood cancer manually is possible by a lumbar puncture, lymph node biopsy, and bone marrow biopsy. White blood cancer cells can be misdiagnosed and the patient's life put at risk if they are detected manually. As a result, faster detection and fewer misdiagnoses are achieved by automatic detection. Detecting blood cancer using existing algorithms and approaches has been the focus of this paper.
Keywords: *white blood cell, platelets, red blood cells, leukemia, myeloma, lymphoma, machine learning.*

1. Introduction

A biomedical research point is the computerized analysis of white platelet tumour diseases like Leukemia and Myeloma. A patient's health can be greatly improved by detecting problems early. In the event of a cancer discovery, blood smear images are a reliable and accurate source of information. It is imperative that feature extraction and reduction be done correctly when evaluating cancer growth finding such leukaemia cancer image. Diagnosing blood cancer manually is possible by a lumbar puncture, lymph node biopsy, and bone marrow biopsy. White blood cancer cells can be misdiagnosed and the patient's life put at risk if they are detected manually. As a result, faster detection and fewer misdiagnoses are achieved by automatic detection. Detecting blood cancer using existing algorithms and approaches has been the focus of this paper.

The primary goal is to develop an automatic classification system for microscopic blood pictures from cancer and non-cancer patients. Leukemia, lymphoma, and myeloma are all forms of blood cancer. Leukemia and myeloma, which are both cancers of the white blood cells, can be studied using computerized white platelet tumour analysis.

An estimated 8.8 million people died in 2015, making it the second largest cause of death in the world (behind heart disease), according to the World Health Organization (WHO) [1. The length of time it takes to treat certain disorders can be greatly reduced if they are discovered early. Doctors are also baffled by some of the disease's subtypes. Diagnostic pathology is increasingly relying on computerized methods to assist in the diagnosis [2]. Leukemia and Myeloma, two of the most common forms of white blood cell cancer, pose a serious threat to the lives of many people. To diagnose leukaemia, doctors look for abnormal white blood cells in the bone marrow. It could be short-term or long-term. Acute myeloid leukaemia (AML), acute lymphoblastic leukaemia (ALL), chronic myeloid leukaemia (CML), and chronic lymphocytic leukaemia (CLL) are the four most common kinds of leukaemia (CLL). Anemia is a subcategory of Acute Myeloid Leukemia (AML) (M0, M1, M2, M3, M4, M5, M6, M7). AcuteThe subtypes of lymphoblastic leukaemia include: (L1, L2, L3). Myelomonocytic (M3), Promyelocytic (M3), Myelomonocytic (M4), Monocytic (M5), Erythroleukemia (M6), and Megakaryocytic (M7) are all kinds of acute myelogenous leukaemia (AML) that have undergone specific examination. The symptoms of ALL and AML are so similar to those of other diseases that a proper diagnosis is nearly impossible. A magnifying tool is used for blood tests as part of the leukaemia detection process. Plasma cells, a type of blood white platelet that creates antibodies to help fight disease, are affected by myeloma. B lymphocytes in myeloma don't turn into plasma cells, yet they keep reproducing at an alarming rate. They clog up the bone marrow and prevent it from producing a variety of platelets. – The synthetic components that it expands also break down and remodel bone, causing pain and fissures from time to time. Because the loss of bone marrow can occur anywhere in the body, myeloma is often referred to as "different myeloma." Blood proteins beta2 microglobin and albumin can be used to categories myelomas into three distinct stages. Platelet, red, and white cells are all present in healthy blood cells seen in Figure 1. Figure 2 depicts a malignant cell with a massive amount of white blood cells and blasts.

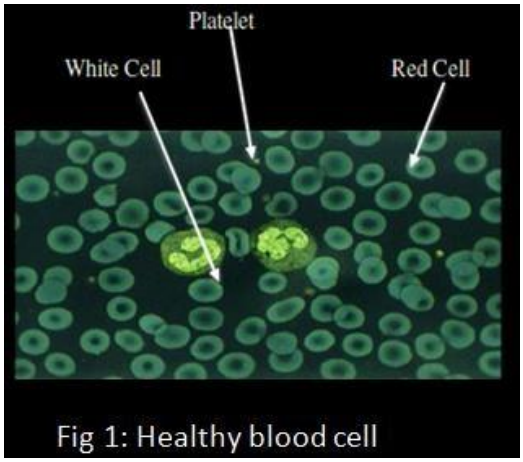


Fig 1: Healthy blood cell

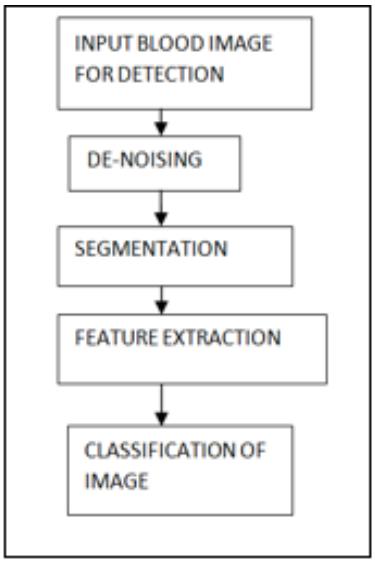
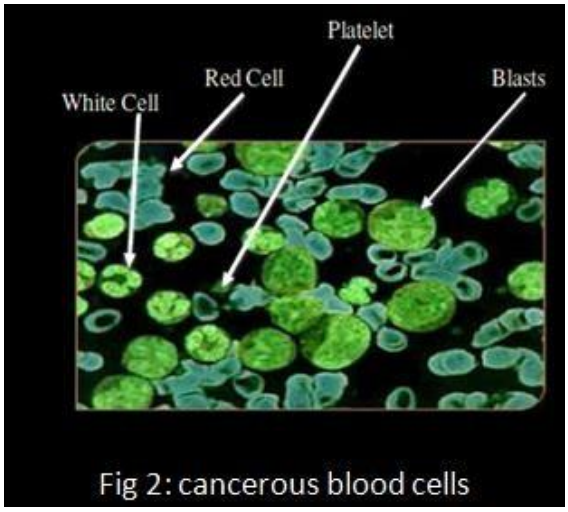


Fig -1: Block diagram of the proposed system



On the basis of factors such as: sensitivity, precision, severity, efficiency, type, and accuracy, we compare some of the methodologies for automatic detection of white blood cancer cells in this study. The purpose of this paper is to review the most relevant work in the subject of automated detection of white blood tumour cells and to provide an in-depth look at the available approaches. Here's how the paper is laid out. Detection of automated white blood cancer cells is described in Section 2. According on the literature review, available methods are presented in section 3. Section 4 explains the accuracy methodologies, followed by Section 5's conclusion.

1. Automated White Blood Cancer Cell Detection General Workflow

White blood cell cancer detection relies heavily on image pre-processing, feature extraction, and classification. Another important component that contributes to the accuracy of the final outcome is the elimination of noise. Automated detection of white blood cancer cells utilising machine learning and image processing techniques is illustrated in the following flowchart..

1.1 Input Image

Image acquisition is a common term for this step. Input images from a microscope are required for the detection of blood cancer disorders such as leukaemia and myeloma. Images of blood smear are included in this informational index or data set for the purpose of assessing and comparing algorithms for segmentation and image classification.

1.2 Image Pre-Processing

The supplied image should be de-noised as much as possible. Color relevance wheel is also included in pre-processing. Grey colour space images are created from RGB photos..

1.3 Segmentation Of The Image

It's the technique of dividing a large image into multiple smaller pieces. Cancer cell detection accuracy is strongly influenced by this key stage..

1.4 Feature Extraction

This is a technique for reducing the dimensionality of data by turning it into a series of high points. Extraction of key features will help to separate important data from irrelevant input. It has features like Haar wavelet ones; Hausdorff ones with and without LBP; Shape ones; Gray level co-occurrence ones; Haralick ones or GLCM; Color one. It also has features like these:

3. Classification

Classification is the final step in the process. An picture of blood that includes a white blood cancer cell can be detected using this technique. Random forests, support vector machines, closest neighbour algorithms, neural network techniques, and decision tree algorithms are all included.

Existing methods in automated cancer cell detection technologies: a literature review

The use of automated detection in the detection of white blood cancer cells is highly recommended. Oncologists' work can be more accurate thanks to an automated method, and the expense of detecting cancer can also be cut. It is possible to drastically minimise the rate of misdiagnosis and obtain early detection of cancer. Automated cancer cell detection now relies heavily on image processing, segmentation, feature extraction, and classification. Acquisition, improvisation, Representation/ Transformation/Restoration/ Transformation/ Restoration/ Processing of Color Images/Compression/Image Processing Morphological Image Processing Segmentation/Object Recognition.

This article discusses how image processing can be used to detect leukaemia. There are generally three stages to the strategy for identifying leukaemia in blood tests using microscopic images:

Images of cells are segmented and features are extracted from them. Then, the malignant development cells are classified.

Researchers Hend Mohamed, Rowan Omar, Nermeen Saeed, Ali Essam, Nada Ayman, Taraggy Mohiy, and Ashraf AbdelRaouf [1] devised a hybrid method for detecting white blood cancer cells that combined deep learning with support from a learning system. [2] It was utilised for segmentation and a random forest as a

classifier, as shown in Figure 1.

This method achieved an accuracy rate of 93 percent for L1, L2, M5, M3 and Myeloma based on a total of 105 pictures.

Methodology developed by Anitta K Varghese et al. [4] is simple and effective in classifying full blood smear images.

To identify and segment nucleated cells. K-mean division and layer subtraction division are used to complete the division.

Dimensionality was reduced by using Hausdorff measurements. Direct SVM two-class classification is used because it is inexpensive and provides a good level of performance.

Niranjan Chatap and Sini Shibu [5] devised an iterative Threshold computation for the purpose of dividing noisy images. This formula solves the problem of extracting and dividing cells from large, raucous images. The morphological approach to cell picture division is more precise than the traditional watershed-based method. Blood smear images are used as a starting point, and a basic thresholding strategy to determining the calculation is used.

Reinforcement learning was proposed by W., Qiang, Zhongli, Z., [6]. It divides leukaemia into acute myeloid leukaemia (AML), acute lymphoblastic leukaemia (ALL), chronic myeloid leukaemia (CML) and chronic lymphocytic leukaemia (CLL).

S. Mohapatra et al. [7] in 2010 presented Unsharp Masking Sub-Imaging Bounding Box as selective filtering and applied Fuzzy C-Means Clustering followed by Nearest Neighbor Classification using support vector machine (SVM).

Local binary patterns, the histogram of gradients, and an already-trained deep network were used by Taha J. Alhindi [8] to extract features from images. Support vector machines, decision trees, and an artificial neural network were utilised to classify images.

For testing the accuracy and precision of classification and highlight or feature extraction models used in the examination, KIMIA Path960, a free dataset of 960 histopathological images extricated from 20 distinct tissue scans, is available. SVM, which uses local binary patterns as features, reaches the maximum accuracy of 90:52%, surpassing the 81:14% achieved by.

In the preprocessing stage, Karthikeyan and Poornima [9] introduced an adaptive median filter for noise reduction and an adaptive Histogram Equalization algorithm for contrast enhancement. For segmentation, they used k-means and fuzzy c-means clustering. For categorization, they used Support Vector Machine (SVM). By employing fuzzy logic and k-means, we were able to obtain a 90% success rate.

Another approach [10] to detect acute leukaemia was proposed by Tejashri G. Patil and V. B. Raskar based on Otsu's segmentation algorithm and feature extraction using contour signature.

According to Mu-Chun Su et al. [11] 2014, "A Neural-Network Based Approach to White Blood Cell Classification" integrates HSI Color Space for preprocessing and discriminating region morphological operators and 7x7 median filter for image postprocessing in order to segregate five categories of WBC. SVM, Multilayer Perceptron, and Hyperrectangular Composite Neural Networks are some of the classification algorithms used by classifiers (HRCNN). MLP had a 99.11 percent accuracy rate, followed by svm with 97.55 percent, and HRCNN with 88.95 percent.

Support Vector Machine(SVM) classifier was used by Khot S.T et al[12] to differentiate leukemic cells. Connecting the highlights to a classifier separates them from the rest of the image.

Using image processing and data mining, Carolina Reta, Leopoldo Altamirano, Jesus A. Gonzalez, Ranquel Diaz, and Jose S. Guichard [13] developed a two-phase method for analysing white blood cancer cells. Markov random fields have been presented as a new way for segmenting colour and texture information. CIE For classifying data, decision trees and the L*a*b* colour space were both used.

For the evaluation of their strategy, Sarrafzadeh et al. [14] offered a new methodology based on the separation of several sub-sorts of M2. 27 tiny images of three AML subtypes were used: 9 AML of M2; 10 AML of M3; and 8 AML of M5. Foregrounding and highlighting were accomplished by utilising L*a*b** shading techniques. K-means clustering is used to separate leukocytes from other blood cells. Discriminative Dictionary Learning is used to extract features that can then be categorised (DDL). The Medical Image and Signal Handling Research Centre (MISP) dataset had an accuracy rate of 97.53 percent.

[15] Harun et al. (2015) suggested K-Means, Fuzzy C-Means, Moving K-Means, and a seeded region growing area.When SVM was utilised as an image classifier, it had accuracy of 98.2 percent.To identify white platelets using magnifying instrument photographs,

Piuri et al[16] focused on this topic in their paper. All of the leucocytes can be discovered and ordered by using a magnifying lens shading picture that recognises the accompanying classes of cells.Eosinophilia, Lymphocytosis, Monocyteism, and Neutrophilism. When a blood image is taken, a neural classifier is used to classify leucocytes based on their physical characteristics, as well as their location in the blood.

Singa et al. (2002 [17]) devised a two-part strategy for WBC identification in colour photos utilising the HSV colour space, K-means, EM algorithm that yielded an accuracy of 80 percent.

Shape-based characteristics were presented by Himali P. Vaghela et al[18] and are more precise than watershed transform, K means clustering, and histogram equalising methods for counting leukemic cells and checking leukemic cells, with a precision of 97.80%.

Bhattacharjee and Saini [19] have presented a watershed transformation technique-based approach to detect ALL. Prior to division, they linked picture improvement and quality alteration.

They used a watershed computation to separate the platelet core from the cell core during division. Binary Search Tree (BST) classified 86% of the data, whereas Gaussian Mixture Models (GMM) classified 93% of the data. Blood smears can be identified and segmented using a technique provided by Agaian et al. [20]. K-Means clustering was used to segment the CIELAB Color space.

In order to compute the Hausdorff Dimension features, the box counting method and the Local Binary Pattern were employed (LBP). Classification accuracy of 98 percent was attained using the Support Vector Machine (SVM).

4. Measuring performance

These methods and measures are discussed in the next section. S. AGAIAN ET AL. (2014), in order to ensure the suitability of our framework, evaluated the performance using specific measures (value parameters). Precision, specificity, sensitivity, F-measure, and accuracy are all examples of these measurements. Trying to classify a specimen results in four possible outcomes: True Positives (TP), where cancer cells are correctly distinguished; False Positives (FP), where benign cells are incorrectly identified as dangerous; True Negatives (TN), where benign cells are correctly identified as dangerous; and False Negatives (FN), where malignancy cells are incorrectly identified as noncancerous. Table I shows that greater estimates of these factors lead to better precision because of the general exactness of the tested image that is given in detail.

Table -1: Performance evaluation parameters

PARAMETERS	DEFINITION	FORMULAE
Sensitivity (recall)	Test's ability to identify positive results.	Se=(TP/(TP+FN))
Precision	proportion of cases with positive Results, which are correctly identified.	P=(TP/(TP+FP))
Specificity	Test's ability to identify negative results.	Sp=(TN/(TN+FP))
F-measure(overal classification performance)	It is the harmonic mean of precision and Sensitivity.	F=(2*Sp*P)/(P+Se)
Accuracy	Measure of the degree of closeness of a measured or **calculated** value to its actual value.	A=((Tp+Tn)/(Tp+Tn+Fp+Fn))

5. Conclusion

This study quickly summarises and analyses relevant work on automated blood cancer cell detection using machine learning and image processing approaches. Leukemia detection can be sped up with the help of image processing. The administrator's magnifying glass can be defeated by an image processing approach, especially if the inquiry was influenced by the amount of interest and health of the observer. It is difficult to detect white blood cancer cells using an automated system. At the earliest opportunity, the automated system should be able to correctly detect the intended goal. Support vector machine (SVM) classifiers and k-means segmentation algorithms are used in the majority of approaches. Segmentation algorithm, feature extraction method, and classification method all play a significant role in determining the accuracy of an automated system.

References

1. Hend Mohamed, Rowan Omar, Nermeen Saeed, Ali Essam, Nada Ayman, Taraggy Mohiy and Ashraf AbdelRaouf "Automated Detection of White Blood Cells Cancer Diseases", in IEEE 2018 First International Workshop on Deep and Representation Learning (IEDRL).
2. J. B. Henry, J. P. AuBuchon, Clinical diagnosis and management by laboratory methods, Archives of Pathology and Laboratory Medicine 121 (9) (1997) 1016.
3. K. M. Sensing, Cancer. URL http://www.who.int/mediacentre/factsheets/fs 297/en.html
4. Anitta K Varghese and Nisha J S,"Automated Screening System for Acute Myelogenous Leukemia Detection using Layer Subtraction",in:International Journal of Current Engineering and Technology ,Available online 02 Oct 2015, Vol.5, No.5 (Oct 2015)
5. Niranjan Chatap, Sini Shibu" Analysis of blood samples for counting leukemia cells using Support vector machine and nearest neighbour" (IOSR-JCE) e-ISSN: 2278-0661,p-ISSN: 2278- 8727, Volume 16, Issue 5, Ver. III (Sep – Oct.2014), PP 79-87.
6. W., Qiang, Zhongli, Z., 2011. "Reinforcement Learning,Algorithms and Its Application", International Conference on Mechatronic Science, Electric Engineering and Computer,,Jilin, China, August 19-22, 2011, pp. 143-1146.
7. S. Mohapatra, S. S. Samanta, D. Patra, S. Satpathi, Fuzzy based blood image segmentation for automated leukemia detection, in: Devices and Communications (ICDeCom), 2011 International Conference on, IEEE, 2011, pp. 1–5.
8. Taha J. Alhindi, Shivam Kalra, Ka Hin Ng, Anika Afrin, Hamid R. Tizhoosh ,"comparing LBP,HOG and deep features for classification of histopathology images", in: IEEE world congress on computional intelligence (IEEE WCCI),2018.
9. T. Karthikeyan, N. Poornima, Microscopic image segmentation using fuzzy c means for leukemia diagnosis, Leukemia 4 (1).
10. Tejashri G. Patil, V. B. Raskar.," Blood microscopic image segmentation & acute leukemia detection"(IJERMT) ISSN: 2278-9359 (Volume-4, Issue-9)SEP 2015.
11. M. Su, C. Cheng, P. Wang. "A Neural-Network-Based Approach to White Blood Cell Classification." The Scientific World Journal Volume 2014, Article ID 796371. 2014.
12. Khot S.T, Sneha Bhalekar, Divya Jaggi and Dolly Rani., "An innovative approach in myelogenous leukemia detection using attribute analysis" (IJAEEE) ISSN (Print): 2278-8948, Volume-2, Issue-5, 2013.
13. Carolina Reta, Leopoldo Altamirano, Jesus A. Gonzalez, Ranquel Diaz and Jose S. Guichard (2015), " segmentation of bone marrow cells images for morphological classification of acute leukemias". Twenty-Third International Florida Artificial Intelligence Research Society Conference (FLAIRS 2010).

14. O. Sarrafzadeh, H. Rabbani, A. M. Dehnavi, A. Talebi, Detecting different sub-types of acute myelogenous leukemia using dictionary learning and sparse representation, in: Image Processing (ICIP), 2015 IEEE International Conference on, IEEE, 2015, pp. 3339–3343.
15. N.H. Harun, Nasir, A.S. Abdul, Mashor, M.Y. Mashor, R. Hassan, "Unsupervised Segmentation Technique for Acute Leukemia Cells Using Clustering Algorithms." International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol:9, No:1, 2015.
16. V. Piuri, F. Scotti: "Morphological Classification of Blood Leucocytes by Microscope Images" Proc. International Symposium on Computational Intelligence for Measurement Systems and Applications, Boston, MD, USA, July 2004.
17. N.Sinha, A.G.Ramakrishnan,"automation of differential blood count",2003
18. Himali P. Vaghela, Hardik Modi, Manoj Pandya, Manoj Pandya.," Leukemia Detection using Digital Image Processing Techniques" (IJAIS) – ISSN : 2249-0868 Volume 10 – No.1, November 2015
19. R. Bhattacharjee, L. M. Saini, Detection of acute lymphoblastic leukemia using watershed transformation technique, in: Signal Processing, Computing and Control (ISPCC), 2015 International Conference on, IEEE, 2015, pp. 383–386.
20. S. Agaian, M. Madhukar, A. T. Chronopoulos, Automated screening system for acute myelogenous leukemia detection in blood microscopic images, IEEE Systems journal 8 (3) (2014) 995–1004.
21. D.Goutam and S.Sailaja ," Classification of Acute Myelogenous Leukemia in Blood Microscopic Images using Supervised Classifier", in: 2015 IEEE International Conference on Engineering and Technology (ICETECH), 20th March 2015, Coimbatore, TN, India.
22. http://www.cancercenter.com/leukemia/types
23. [ALL-IDB] Acute Lymphoblastic Leukemia Image Database for Image Processing. Department of Information Technology – Universita degli Studi di Milano. [cited 2016 Nov
24. https://www.lls.org/leukemia
25. Medical image and signal processing research center. URL http://misp.mui.ac.ir/en
26. I. Steinwart, A. Christmann, Support Vector Machines, Information Science and Statistics, Springer New York, 2008.