RESEARCH ARTICLE OPEN ACCESS

Automated Detection of White Blood Cells Cancer Diseases

Mrs R Jhansi Rani MCA [1], Koduru Poojitha [2]

[1] Asst. Professor, Department Computer Applications
[2] Student, Department of Computer Applications
[1], [2] Chadalawada Ramanamma Engineering College (Autonomous)

ABSTRACT

Mechanized analysis of white platelets malignant growth infections, for example, Leukemia and Myeloma is a difficult biomed-ical inquire about point. Our methodology introduces out of the blue another best in class application that helps with diagnosing the white platelets infections. we break these sicknesses into two classifications, every classification contains like side effects infections that may confound in diagnosing. In light of the specialist's determination, one of two methodologies is actualized. Each methodology is connected on one of the two maladies classification by processing distinctive highlights. At last, Random Forest classifier is connected for ultimate choice. The proposed methodology means to early disclosure of white platelets malignancy, decrease the misdiagnosis cases notwithstanding improve the framework learning approach. In addition, permitting the specialists just to have the last tuning on the outcome acquired from the framework. The proposed methodology accomplished an exactness of 93% in the principal classification and 95% in the second class.

Keywords: - Mechanized analysis of white platelets malignant growth infections, for example, Leukemia and Myeloma is a difficult biomed-ical inquire about point

I. INTRODUCTION

The blood usually consists of cells that are special, called as plasma. Blood is made up of 55% of plasma and 45% of formed elements. A lot of important functions are performed by the blood. The erythrocytes contain hemoglobin which helps to carry oxygen to the tissues and also collect the carbon-dioxide. The blood also has other nutritive substances such as mineral slats, sugar, amino acid, etc. Generally, the more dreadful the disease is, the more time is taken for the diagnosis. Due to the patients waiting time the chance of death is highly increased. Hence it is necessary to diagnose a disease in less time with more accuracy. The time taken for the results to come is only within a week. Diseases such as Ebola, Hepatitis and Corona take a lot time for detection. While these diseases should be treated immediately. And also, pathological tests are expensive and are sometimes not affordable to the patients. Hence the blood reports must be detected automatically in less time and should be cost effective.

The main purpose is to develop an automated classification system, which segments and classifies microscopic blood images from cancer and non-cancer people. The three types of blood cancer are leukemia, lymphoma and myeloma. Computerized analysis of white platelets tumor ailment such as, Leukemia and Myeloma is a testing biomedical look into point. According to the World Health Organization (WHO) [1], Cancer is considering as the second leading cause of death in the world and has claimed the lives of 8.8 million people in 2015, attributed to the death of nearly one in six deaths worldwide. Discovering those diseases in the early stages highly affects the treatment period. In addition, some of the diseases' sub-types are really confusing to the Doctors. Nowadays, there is a great tendency for diagnostic pathology to heavily rely on automated systems which can aid in the diagnosis [2].

White blood cells cancer diseases; Leukemia and Myeloma, threaten people's life nowadays. Leukemia is found when the bone marrow produces abnormal white blood cells, which don't function properly. It may be either acute or chronic. There are four main types of leukaemia: Acute myeloid leukaemia (AML), Acute lymphoblastic leukaemia (ALL), Chronic myeloid leukaemia (CML) and Chronic lymphocytic leukaemia (CLL). Acute Myeloid Leukemia (AML) is sub categorized to (M0, M1, M2, M3, M4, M5, M6, M7). Acute

lymphoblastic leukemia (ALL) is sub categorized to (L1, L2, L3). The types of acute myelogenous leukemia included: Myeloblastic (M0) on special analysis, Myeloblastic (M1) without maturation, Myeloblastic (M2) with maturation, Promyelocytic (M3), Myelomonocytic (M4), Monocytic (M5), Erythroleukemia (M6), Megakaryocytic (M7) [23]. The Symptoms of ailment ALL and AML usually like different sicknesses indications, so the analysis is very troublesome.

One phase of leukemia determination systems concerns examinations in the blood utilizing a magnifying instrument. Myeloma influences plasma cells, a kind of develop white platelet in blood that produces antibodies to enable battle to off diseases. In myeloma, B lymphocytes don't develop legitimately into plasma cells yet keep on repeating quickly. These youthful cells aggregate in bone marrow and keep it from delivering different sorts of platelet. It additionally expands synthetic compounds that breakdown and redesign bone, prompting torment and cracks at times.

Myeloma is frequently alluded to as different myeloma since this diminishing of bones can happen anyplace in body where there is bone marrow. Myelomas can be gathered into three phases dependent on the level of two proteins (called beta2 microglobin and albumin) in blood.

Figure 1 depicts a healthy blood cell with good quantity of platelet,red and white cell. A cancerous cell with enormous amount of white blood cells and blasts is shown in figure 2.

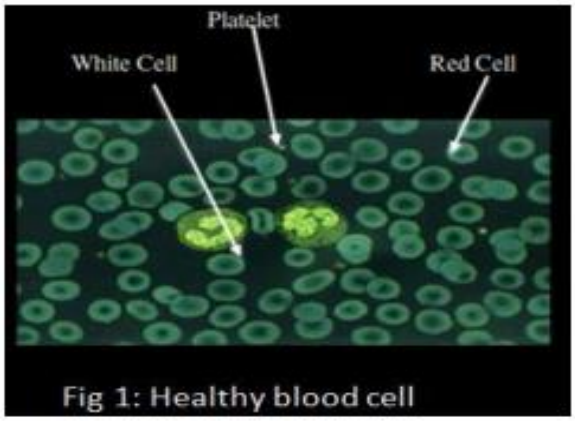

Fig 1: Healthy blood cell

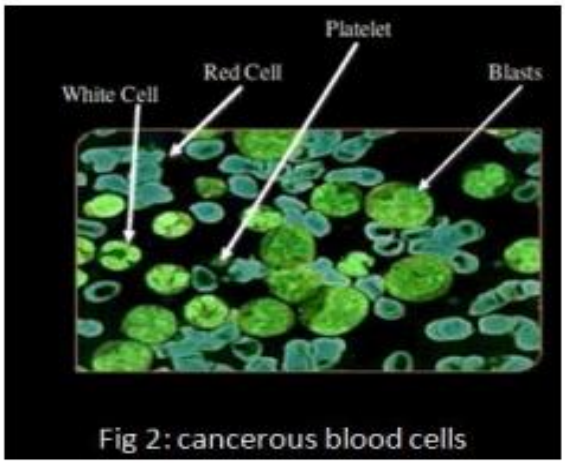

Fig 2: cancerous blood cells

II. RELATED WORK

Some research were proposed to differentiate between: 1) AML and ALL diseases:

1) Karthikeyan and Poornima [11] proposed an approach for detection of Leukemia in blood at early stages. They have used adaptive median filter for noise removal and adaptive Histogram Equalization for contrast enhancement in preprocessing stage. They applied kmeans and Fuzzy c-means clustering for segmentation. They computed statistical, textural and geometrical features and applied Support Vector Machine (SVM) for classification. Their approach achieved 90% with Fuzzy c-means and 83% with k-means using Fuzzy Logic: Intelligence, Control, and Information dataset [12]. Another research by Mohapatra et al. [13] proposed a Fuzzy based Blood Image Segmentation for Automated Leukemia Detection. They applied selective median filtering followed by unsharp masking in preprocessing. In segmentation, they used improved version of fuzzy clustering technique viz. Gustafson Kessel clustering [14] followed by nearest neighbor classification in L*a*b* color space (L* for lightness, a* for rednessgreenness axis, and b* a yellownessblueness axis) [15]. The computed features are two novel shape features; Hausdorff Dimension and contour signature. Support Vector Machine (SVM) is employed for classification and they achieved 93% on a database of 49 blood smear images of size 512 x 512 pixels.

2) AML detection only or ALL detection only: An approach by Agaian et al. [16] proposed a simple technique that automatically detects and segments AML in blood smears. Segmentation was done in the CIELAB Color space by K-Means clustering algorithm. Hausdorff Dimension features were computed using the box counting method and Local Binary Pattern (LBP). Classification achieved 98% accuracy using the Support Vector Machine (SVM) on American Society of Hematology (ASH) for Leukemia dataset [17]. This dataset comprised of 80 images40 from AML patients and 40 from non-AML patients. The image size used for their classification was 184 x 138 pixels. Another system proposed by Bhattacharjee and Saini [18] for Acute Lymphoblastic Leukemia detection using Watershed Transformation Technique. They applied contrast enhancement and quality adjustment for enhancing images before segmentation. In segmentation they used watershed algorithm, isolating the blood cell and the cell nucleus. They computed area, perimeter, circularity and form factor features. Gaussian Mixture Models (GMM) and Binary Search Tree (BST) were applied for classification. GMM achieved 93% while BSTe achieved 86%. They applied their approach on 150 lymphocytic cells images (30 normal cells & 120 blast cells) accessed from ALL-IDB1 & ALLIDB2 datasets [17].

3) Detection of (AML) sub-types: Another proposed approach by Sarrafzadeh et al. [19] focused mainly on differentiate between M2, M3 and M5 sub-types to evaluate their introduced method. The approach was applied in the L*a*b* color space. Segmentation is performed using K-means clustering to segregate leukocytes from other blood components. Texture and shape features are extracted in order to be classified using Discriminative Dictionary Learning (DDL). They achieved 97.53% accuracy on Medical Image & Signal Processing Research Center (MISP) dataset [20]. They used a dataset composed of 27 microscopic images of three sub-types of AML; 9 AML-M2, 10 AML-M3 and 8 AML-M5.

III. EXISTING SYSTEM

- In existing system, an approach for detection of Leukemia in blood at early stages have used adaptive median filter for noise removal and adaptive Histogram Equalization for contrast enhancement in preprocessing stage.
- They applied k means and Fuzzy c-means clustering for segmentation. They computed statistical, textural and geometrical features and applied Support Vector Machine (SVM) for classification.
- Their approach achieved 90% with Fuzzy c-means and 83% with k-means.

Disadvantages

- Less Accuracy
- Time taken process

IV. PROPOSED SYSTEM

- In proposed system, we propose the design, development and evaluation of an automated system to accurately detect white blood cells cancer diseases. It detects types and sub-types of Leukemia (ALL and AML) and Myeloma.
- It is a recognition system applied on acquired blood microscopic images then performs preprocessing, segmentation, feature extraction and classification.
- Using the novel algorithm for detects all the sub types.
- To increase accuracy of detection and classification, use good machine learning models.

Advantages

- The proposed solution converts images to YCBCR color space and construct Gaussian distribution of CB and CR values. Statistical, texture, size ratio and morphological features are then computed to train classifier.

- Unlike existing systems, our system has the ability of learning from misclassified tests to enhance the future accuracy of the system. Random Forest classifier is the best classifier that is able to differentiate between different types and the one which gives us the best accuracy.

V. RGB COLOR IMAGE

The RGB covering model is an extra matter shading classical in which red, green, and blue light are incorporated into various ways to deal with reproduce a wide show of tones. The name of the model begins from the initials of the three included substance fundamental tones, red, green, and blue. The essential inspiration driving the RGB shading model is for the recognizing, depiction, and show of pictures in electronic structures, for instance, TVs and PCs,

anyway it has furthermore been used in customary photography. Preceding the electronic age, the RGB shading model recently had a solid theory behind it, arranged in human perspective on colors.RGB is a device subordinate shading model: assorted contraptions recognize or mimic a given RGB regard in a startling manner, since the shading parts, (for instance, phosphors or hues) and their response to the individual R, G, and B levels change from maker to producer, or even in a comparable device after some time. Thusly a RGB regard does not portray a comparative shading transversely over devices without some kind of shading the board. Run of the mill RGB input strategies are shading TV and camcorders, picture scanners, and automated cameras. Normal RGB yield contraptions are TV sets of various advances (CRT, LCD, plasma, etc.), PC and mobile phone appears, video projectors, multicolor LED exhibits, and immense screens, for instance, Jumbo Tron. Shading printers, on the other hand, are not RGB devices, yet subtractive shading devices .

VI. GRAYSCALE

In taking pictures and figuring, a grayscale or greyscale electronic image is a image in which the estimate of every pixel is a solitary example, that is, it conveys just intensity information. Images of this sort, then called extremely conflicting, are complete solely out of tinted lenses of dim, changing from shady at the weakest control to white at the greatest ashore. Grayscale images are specific from one-piece bi-tonal very conflicting images, which with respects to PC imaging are images with just the two types, shady, and white (likewise named bit level or parallel images). Grayscale images have frequent shades of dim in the central. Grayscale images are additionally called monochromatic, signifying the nearness of just one (mono) shading (chrome). Grayscale images are normally the consequence of approximating the power of bright at every pixel in a private group of the electromagnetic series (for example infrared, obvious bright,ultraviolet, and so on.), and in such suitcases they are monochromatic genuine when just a given reappearance is wedged. Yet in adding they can be combined from a full covering image; understand the part about altering ended to grayscaleExampleofgrayscaleimageisgivenbelow.
VII.

VII. RESULTS

To locate the precise highlights we need to portion the lung district from the chest CT check picture for simple calculation. For fragmenting the lung district from the chest CT examine picture morphological task is completed. We characterized a picture as a (plentifulness) capacity of two, genuine (organize) factors a(x,y) or binary, separate factors a[m,n]. An elective meaning of a image can be founded on the supposed that a image includes of a usual (or gathering) of also nonstop or separate orders. One might say the set links to the attentions or pixels that have a residence with the items trendy the image. This is delineated in number beneath which contains two exercises or sets A and B. Memo that the arrange outline is obligatory. For the minute we will consider the pixel esteems to be paired as examined in the Further we will confine our talk to discrete space.

Preprocessing

The reason for the pre-handling stage is to expel undesirable impacts, for example, clamor from the picture, and change or modify the picture as essential for further preparing. The goals of the picture is

diminished by a factor of four to 512·384 to accelerate execution of the framework. Likewise, the test pictures will be exposed to specific middle separating and unsharp covering to segregate clamor which may have been aggregated amid picture securing and because of over the top recoloring.

Segmentation

The strategy of apportioning the picture into portion can be characterized as picture division. Considering the comparative property, segmentation is executed. This comparative property is bunch together our propounded methodology actualizes Lloyd's clustering technique which helps in the division of blood minute pictures based on alike properties. This method broadens the k-mean bunching calculation by presenting rehashed division conspire which investigates the centroid of each set in the segment and in the long run re-portion the info dependent on the nearest centroid. This method helps in the extraction of important image attributes, in light of which data can be effectively perceived.A basic thresholding method stands connected to stretch beginning names to pixels trendy the platelet pictures.The calculation depends on apriori data about blood smear pictures. At that point the marks are balanced with a shape discovery technique dependent on vast regional context data to deliver significant outcomes.

VIII. CONCLUSION

In this paper we propose the design, development and evaluation of an automated system to accurately detect white blood cells cancer diseases. It detects types and sub-types of Leukemia (ALL and AML) and Myeloma. It is a recognition system applied on acquired blood microscopic images then performs preprocessing, segmentation, feature extraction and classification. The proposed solution converts images to YCBCR color space and construct Gaussian distribution of 54 CB and CR values. Statistical, texture, size ratio and morphological features are then computed to train classifier. Unlike existing systems, our system has the ability of learning from misclassified tests to enhance the future accuracy of the system. Random Forest classifier is the best classifier that is able to differentiate between different types and the one which gives us the best accuracy. The system achieved 94.3 % accuracy in detecting and classifying types and sub-types. As our next step, we aim to detect more types of white blood cells cancer diseases to build an overall system for white blood cells diseases.

REFERENCES

[1] K. M. Sensing, Cancer. URL http://www.who.int/mediacentre/factsheets /fs297/en.html

[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] E. Sarhan, E. Khalifa, A. M. Nabil, Post classification using cellular automata for landsat images in developing countries, in: International Conference on Image Information Processing (ICIIP), Shimla, India, IEEE, 2011, pp. 1–4.

[4] A. AbdelRaouf, C. A. Higgins, T. Pridmore, M. I. Khalil, Arabic character recognition using a haar cascade classifier approach (hcc), Pattern Analysis and Applications 19 (2) (2016) 411–426.

[5] A. AbdelRaouf, C. A. Higgins, T. P. Pridmore, M. I. Khalil, Arabic corpus enhancement using a new lexicon/stemming algorithm., in: 2nd International

Conference on Pattern Recognition Applications and methods (ICPRAM), Barcelona, Spain, 2013, pp. 435–440.

[6] T. M. Ghanem, M. N. Moustafa, H. I. Shahein, Gabor wavelet based automatic coin classsification, in: 16th IEEE International Conference on Image Processing (ICIP), Cairo, Egypt, IEEE, 2009, pp. 305–308.

[7] S. Sameh, M. A. Azim, A. AbdelRaouf, Narrowed coronary artery detection and classification using angiographic scans, in: 12th International Conference on Computer Engineering and Systems (ICCES), Cairo, Egypt, IEEE, 2017, pp. 73–79.

[8] S. Moataz, M. A. Azim, A. AbdelRaouf, Automated fish signals fusion detection for chronic myeloid leukemia diagnosis, in: 26th International Conference on Computer Theory and Applications (ICCTA), Alexandria, Egypt, IEEE, 2016.

[9] M. F. for Medical Education, Research, Leukemia. URL https://www.mayoclinic.org/diseases-conditions /leukemia/symptoms-causes/syc-20374373.html [10] What is multiple myeloma? URL https://www.cancer.org/cancer/multiple-myeloma/ about/what-is-multiple-myeloma.html

[11] T. Karthikeyan, N. Poornima, Microscopic image segmentation using fuzzy c means for leukemia diagnosis, Leukemia 4 (1).

[12] J. Yen, R. Langari, Fuzzy logic: intelligence, control, and information, Vol. 1, Prentice Hall Upper Saddle River, NJ, 1999.

[13] 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.

[14] D. Graves, W. Pedrycz, Fuzzy c-means, gustafson-kessel fcm, and kernel-based fcm: A comparative study, in: Analysis and Design of Intelligent Systems using Soft Computing Techniques, Springer, 2007, pp. 140–149.

[15] K. M. Sensing, Identifying color differences using l*a*b* or l*c*h* coordinates. URL https://sensing.konicaminolta.us/blog /identifying-color-differences-using -l-a-b-or-l-c-h-coordinates/

[16] 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.

[17] R. D. Labati, V. Piuri, F. Scotti, All-idb: The acute lymphoblastic leukemia image database for image processing, in: Image processing (ICIP), 2011 18th IEEE international conference on, IEEE, 2011, pp. 2045–2048.

[18] 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.

[19] 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.

[20] Medical image and signal processing research center. URL http://misp.mui.ac.ir/en

[21] T. Acharya, P.-S. Tsai, JPEG2000 standard for image compression concepts, algorithms and VLSI architectures, Wiley-Interscience, 2005.