

Leukemia Detection using Object Oriented Method

Dr. M. Rama Bai
Computer Science Engineering
Mahatma Gandhi Institute of Technology
Hyderabad, India

Sai Sreeja Alapati
Computer Science Engineering
Mahatma Gandhi Institute of Technology
Hyderabad, India

Abstract—Blood cancers affect the production and function of your blood cells. In most blood cancers, the normal blood cell development process is interrupted by uncontrolled growth of an abnormal type of blood cell. This paper will present an approach for the detection and tracking of Leukemia. This paper aims at identifying leukemia, by using image segmentation. A microscopic image of blood sample is taken, the cell count is extracted using image processing module by module. The purpose is to increase the accuracy using three different approaches object wise.

I. INTRODUCTION

This paper aims at identifying leukemia, by using image segmentation. An image is a way of transferring information, and the image contains lots of useful information. Understanding the image and extracting information from the image to accomplish works is an important area of application in digital image technology, and the first step in understanding the image is the image segmentation.

Digital images were acquired using a digital camera connected to a light microscope. In practice, it is often not interested in all parts of the image, but only for some certain areas which have the same characteristics. Image segmentation is one of the hot spots in image processing and computer vision. It is also an important basis for image recognition. It is based on certain criteria to divide an input image into a number of the same nature of the category in order to extract the area which people are interested in. And it is the basis for image analysis and understanding of image feature extraction and recognition.

In this paper we use region-based segmentation algorithm. Region-based segmentation algorithms operate iterative by grouping together pixels which are neighbors and have similar values and splitting groups of pixels which are dissimilar in value. The advantage of regional based growth is that it usually separates the connected regions with the same characteristics and provides good boundary information and segmentation results. The idea of regional based growth is simple and requires only a few seed points to complete. And the growth criteria in the growing process can be freely specified. Finally, it can pick multiple criteria at the same time.

Blood Cancer is a disorder that affects around 1.2 million people worldwide every year. Blood cancer, or hematologic cancers, affect the production and function of blood cells. Most of these cancers start in the bone marrow where blood is produced. Their symptoms usually

come on slowly, so you might not even notice them. And some people have no symptoms at all.

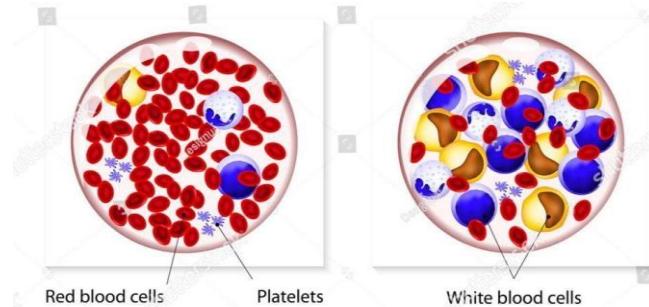


Fig. 1. Normal vs. Effected Cells

Referring to the images in Fig 1, the left image shows us the composition of blood cells of a normal human being with a high number of red blood cells whereas the right one is of a cancer affected human with high composition of white blood cells.

The three most common types of Leukemia are Chronic lymphocyte, Hairy cell, Juvenile Myelomonocytic, Plasma cell, T-cell prolymphocytic.

Chronic Lymphocytic (CLL), is the most common leukemia in adults. It's a type of cancer that starts in cells that become certain white blood cells (called lymphocytes) in the bone marrow. The cancer (leukemia) cells start in the bone marrow but then go into the blood. In CLL, the leukemia cells often build up slowly. Many people don't have any symptoms for at least a few years. But over time, the cells grow and spread to other parts of the body, including the lymph nodes, liver, and spleen.

Hairy Cell Leukemia is a rare, slow-growing cancer of the blood in which your bone marrow makes too many B cells (lymphocytes), a type of white blood cell that fights infection. These excess B cells are abnormal and look "hairy" under a microscope. As the number of leukemia cells increases, fewer healthy white blood cells, red blood cells and platelets are produced. Hairy cell leukemia affects more men than women, and it occurs most commonly in middle-aged or older adults.

Juvenile Myelomonocytic (JMML) is a rare childhood cancer that usually happens in children younger than 2 years old. In JMML, too many myelocytes and monocytes (two types of WBCs) are produced from immature blood stem cells called blasts. These

myelocytes, monocytes, and blasts overwhelm the normal cells in the bone marrow and other organs, causing the symptoms of JMML.

Plasma Cell leukemia (PCL) is an aggressive form of multiple myeloma characterized by high levels of abnormal plasma cells circulating in the peripheral (circulating) blood. Normal plasma cells in the bone marrow produce antibodies that fight infection.

T-cell Prolymphocytic Leukemia (T-PLL) is an extremely rare and typically aggressive malignancy (cancer) that is characterized by the out of control growth of mature T-cells (T-lymphocytes). T-cells are a type of white blood cell that protects the body from infections. T-PLL affects older adults with a median age at diagnosis of 61 years, and it is more common in men than in women.

II. LITERATURE SURVEY

A. AUTOMATED LEUKEMIA DETECTION USING MICROSCOPIC IMAGES:

In this paper, the automated Leukemia detection system analyses the microscopic image and overcomes these draw-backs. It extracts the required parts of the images and applies some filtering techniques. K-mean clustering approach is used for white blood cells detection. The histogram equalization and Zack algorithm are applied for grouping white blood cells. Some of the features like mean, standard deviation, color, area, perimeter, etc. are calculated for leukemia detection. The SVM is used for classification.

B. IMAGE SEGMENTATION USING LOCAL THRESHOLDING AND YCbCr COLOR SPACE:

In this work, the segmentation is performed in YCbCr color space. The segmented image will have 2 different colors, which are black and white, and for this reason the segmentation is done using local threshold values for the Cb component of YCbCr. A mask is used to determine the neighbors of each pixel in the image. The mask also determines an operation to be applied to the neighborhood of every pixel in the image. The mask and the operations are used to determine the local threshold for every pixel in the image. For each pixel location the threshold will be different. This value is compared with the color value of the pixel.

III. MODULES

The paper consists of 5 modules. These 5 modules are:

- Module A: Image module
- Module B: RGB color space module
- Module C: CMYK color space module
- Module D: YCbCr color space module
- Module E: Result calculation module

Fig 2 clearly depicts the modules of our paper where microscopic image is given as input, it processes through the RGB module where it gives the details of red blood cells and white blood cells differently, then the result is calculated by the percentage difference in the count of RBC and WBC , that is how the result is declared by

concluding if leukemia exists or it is normal. The picture is next sent to the CMYK module and later to the YCbCr module where the same process repeats. The output can be exported if needed after the process and saved.

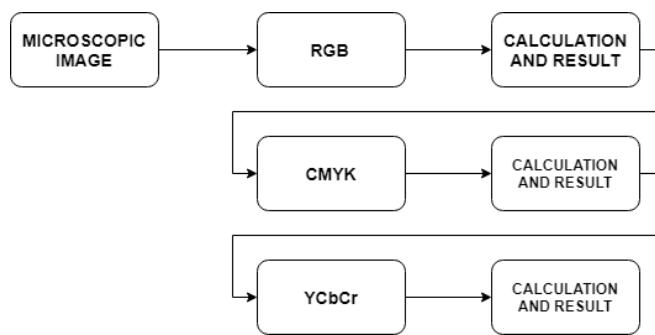


Fig. 2. Architecture diagram

Fig 2 clearly depicts the modules of our paper where microscopic image is given as input, it processes through the RGB module where it gives the details of red blood cells and white blood cells differently, then the result is calculated by the percentage difference in the count of RBC and WBC , that is how the result is declared by concluding if leukemia exists or it is normal. The picture is next sent to the CMYK module and later to the YCbCr module where the same process repeats. The output can be exported if needed after the process and saved.

A. IMAGE MODULE:

A microscopic image of blood smears is taken from ALL-IDB database and given as input to our system through user interface, the system reads the image and then forwards it to the next module.

B. RGB COLOR SPACE MODULE:

An RGB color space is any additive color space based on the RGB color model. RGB color space is defined by the three chromaticities of the red, green, and blue additive primaries, and can produce any chromaticity that is the triangle defined by those primary colors. The complete specification of an RGB color space also requires a white point chromaticity and a gamma correction curve. As of 2007, sRGB is by far the most used RGB color space. RGB is an abbreviation for red green blue.

RGB is a convenient color model for computer graphics because the human visual system works in a way that is similar – though not quite identical – to an RGB color space.

The most commonly used RGB color spaces are sRGB and Adobe RGB (which has a significantly larger gamut). Adobe has recently developed another color space called Adobe Wide Gamut RGB, which is even larger, in detriment to gamut density.

C. CMYK COLOR SPACE MODULE:

The CMYK color model (process color, four color) is a subtractive color model, used in color printing, and is

also used to describe the printing process itself. CMYK refers to the four inks used in some color printing: cyan, magenta, yellow, and key. The CMYK model works by partially or entirely masking colors on a lighter, usually white, background. The ink reduces the light that would otherwise be reflected. Such a model is called subtractive because inks "subtract" the colors red, green and blue from white light. White light minus red leaves cyan, white light minus green leaves magenta, and white light minus blue leaves yellow.

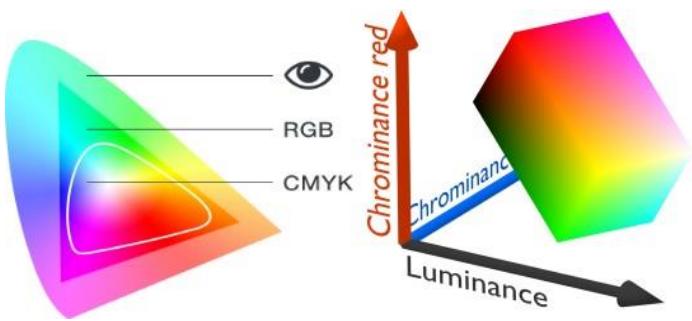


Fig. 3. Color space of RGB and CMYK along with YCbCr color space.

Figure 3 shows the RGB and CMYK in a broad view in visible spectrum on the left graph of color space where CMYK is clearly seemed as the base color perceptive primaries and then we have RGB in the second space where as the visible spectrum is far the next zone concluding that the color we see is not the actual color of the subject. Coming to YCbCr on the right side, it is the color space of digital photography and videography which can almost display all the colors.

D. YCbCr COLOR SPACE MODULE:

YCbCr, also written as $Y'CbCr$ or $Y'CbCr$, is a family of color spaces used as a part of the color image pipeline in video and digital photography systems. Y' is the luma component and C_b and C_r are the blue-difference and red-difference chroma components. Y' (with prime) is distinguished from Y , which is luminance, meaning that light intensity is nonlinearly encoded based on gamma corrected RGB primaries. $Y'CbCr$ color spaces are defined by a mathematical coordinate transformation from an associated RGB color space. If the underlying RGB color space is absolute, the $Y'CbCr$ color space is an absolute color space as well; conversely, if the RGB space is ill-defined, so is $Y'CbCr$.

Y' stands for the luma component (the brightness) and U and V are the chrominance (color) components; luminance is denoted by Y and luma by Y' – the prime symbols ('') denote gamma compression, with "luminance" meaning physical linear-space brightness, while "luma" is (nonlinear) perceptual brightness.

Pixel Based Segmentation:

Image segmentation is the division of an image into regions or categories, which correspond to different

objects or parts of objects. Every pixel in an image is allocated to one of a number of these categories. A good segmentation is typically one in which:

1. Pixels in the same category have similar greyscale or multivariate values and form a connected region.
2. Neighboring pixels which are in different categories have similar values.

E. RESULT CALCULATION MODULE:

From the previous color space modules, we get the count of red blood cells and white blood cells individually. Then the percentage of each of them correspondingly are calculated to the total number of blood cells detected in that particular image.

$$RBC\% = RBC \div (\text{total cells})$$

$$WBC\% = WBC \div (\text{total cells})$$

After the percentage calculation, if the difference in the percentage with WBC greater than 20 is taken as potential leukemia, else declared normal.

$$WBC\% - RBC\% \leq 20$$

This way calculation is done to track and conclude Leukemia occurrence from a microscopic image of blood cells.

Export Data:

The data after the process is successfully done, we get the output on our user interface and we would lose it if we try another blood sample, so we have an "Export" button with which we can save the figure with the data set we obtained as result and continue working with further files without any necessity of noting down the values or without any efforts of taking a snapshot.

IV. RESULTS

The original picture we used for execution is the sample image from ALL-IDB database showing plasma cell leukemia and its execution in three modules. Any type of cancer from the before mentioned categories can be tested in the following manner respectively. No preprocessing of image is required if they are obtained from ALL-IDB database, else they need to be altered with sufficient exposure and highlights for better results.

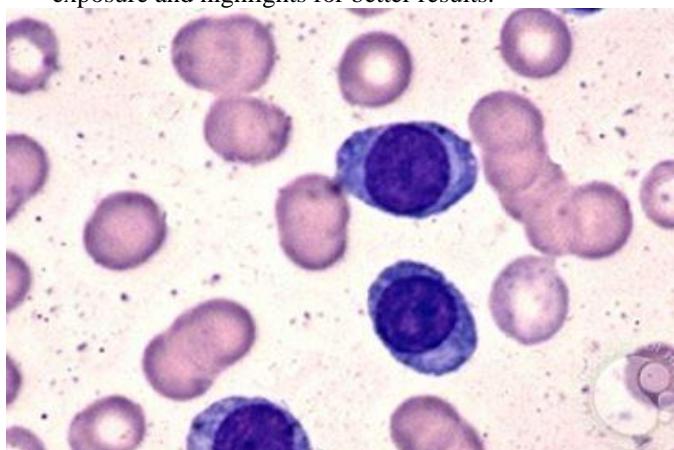


Fig. 4. Plasma Cell Leukemia Sample

For the results to be checked, first the code is executed where we acquire all the possible options. The procedure follows as shown in the figures.

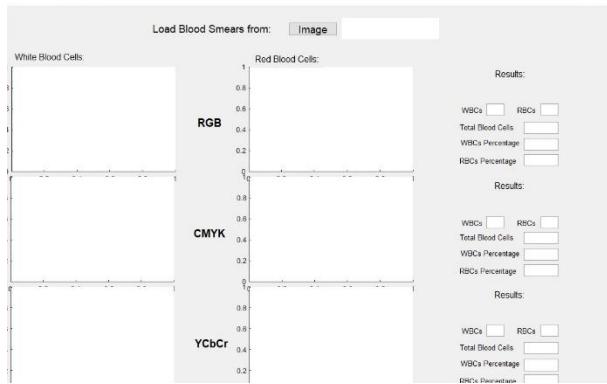


Fig. 5. Output User Interface Screen

Fig 5 is the menu and if we choose our input image from the system and upload it to the system we implemented.

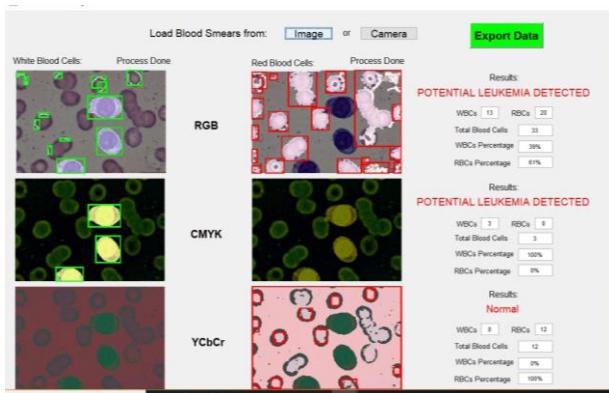


Fig. 6. Output obtained after image processing

Fig 6 shows the execution clearly, RGB, CMYK and YCbCr modules respectively are processed, the result can be concluded by the majority considering all the three-color spaces.

V. CONCLUSION

The paper has presented an approach for the detection and tracking of Leukemia. The proposed system will be able to segment the given microscopic picture in three different approaches to increase the accuracy. With the cell count obtained from both Red Blood and White Blood Cells, we calculate the possibility of Blood Cancer. This method of detection stands strong as the model is tested for all five highly occurring cancers and the testing can be digitalized. This is effective and the cost is absolutely optimized compared to the traditional blood tests.

REFERENCES

- [1] Nimesh Patel,Ashutosh Mishra.Automated Leukemia Detection Using Microscopic images. Sci.2015(12-16).
- [2] Amit Kumar Mandal, Dilip Kumar Baruah.Image Segmentation Using Local Thresholding and Ycbcr Colour Space.J Appl Sci2010.
- [3] Sinha N,Ramakrishnan AG.Automation of differential blood count.In Chock-alangam A,editor.Proceedings of the conference on convergent technologies for the Asia pacific region,October 5-17.Taj Residency Bangalore:IEEE Publisher;2003 p.
- [4] Kovalev VA, Grigoriev AY, Ahn H. Robust recognition of white blood cell images. In: Kavanagh ME, Werner B, editors. Proceedings of the 13th international conference on pattern recognition, August 25-29. Vienna, Austria: IEEE Publisher;1996. p.
- [5] Madhloom Ht,Kareem.An automated white blood cell nucleus localisation and segmentation using image arithmetic and automated threshold.J Appl Sci2010.
- [6] Nobuhito Matsushiro. Color image information processing, 14-16 July. Boston,MA, USA: IEEE Publisher; 2004. p.
- [7] Halim NHA, Mashor MY, Hassan R. Automatic blasts counting for acute leukemia based on blood samples. Int J Res Rev Comput Sci 2011; 2(August (4)).
- [8] Mohapatra S, Patra D, Satpathy S. An ensemble classifier system for early diagnosis of acute lymphoblastic leukemia in blood microscopic images.IEEE Publisher; 2008.
- [9] Donida Labati R, Piuri V, Scotti F. ALL-IDB: the acute lymphoblastic leukemia image Database for image processing. In: Macq Benot, Schelkens Peter, editors. Proceedings of the 18th IEEE ICIP international conference on image processing.