International Journal of Computer Science Trends and Technology (IJCST) – Volume 3 Issue 1, Jan-Feb 2015

RESEARCH ARTICLE OPEN ACCESS

Significant Analysis of Leukemic Cells Extraction and Detection Using KNN and Hough Transform Algorithm

Subhan, Ms. Parminder Kaur
Department of Electronics & Communication Engineering (ECE)
SSIET, Derabassi
Chandigarh – India

ABSTRACT

To recognize the morphological nature of cancerous cell by employing image processing algorithms so that they can be used to automate diagnostic techniques, reducing the cost and increasing the accuracy of epithelial biopsy procedures. Morphological is a very powerful tool in image processing, and it is been used to segment and extract the cancerous cell from the background and other cells. The aim of this research is to produce a computer vision system that can detect and estimate the number of cancerous cell in blood sample. Our main objective is to predict cancer cell in blood samples mainly (Acute leukemia). Acute leukemia is a disease of the leukocytes and their precursors. It is characterized by the appearance of immature, abnormal cells in the bone marrow and peripheral blood and frequently in the liver, spleen, lymph nodes, and other parenchymatous organs .The segmentation of the bone marrow aspirate by applying the Hough transform and KNN Algorithm, selection of individual cells, and feature generation on the basis of texture, statistical and geometrical analysis of the cells and analysis the parameters of abnormal cells such as radius, area, perimeter, shape, color of single abnormal cell. With the reference measurement we can detect the abnormal cells.

***Keywords*:-** White Blood Cells, Segmentation, Blood Cells, Leukemia, MATLAB

I. INTRODUCTION

A. Normal Blood Cells

To understand how leukemia affects blood cells, it helps to know about normal blood cells. All blood cells come from blood stem cells [1]. Although some blood stem cells are in the blood, most are in the bone marrow. Blood stem cells produce three kinds of blood cells:

- **Red blood cells:** Red blood cells carry oxygen all over the body.
- **Platelets** (PLAYT-lets): Platelets help form blood clots to slow or stop bleeding.
- **White blood cells:** White blood cells help fight infection.

B. Normal Myeloid and Lymphoid Cells

A blood stem cell can produce both myeloid stem cells and lymphoid stem cells:

- **Myeloid cells:** A myeloid stem cell can produce red blood cells and platelets. Or, it can produce myeloblasts. (A blast is a type of immature cell.) Myeloblasts can produce several types of white blood cells known as granulocytes[9].
- **Lymphoid cells:** A lymphoid stem cell can produce lymphoblasts, which can produce several types of white blood cells that are different from granulocytes.

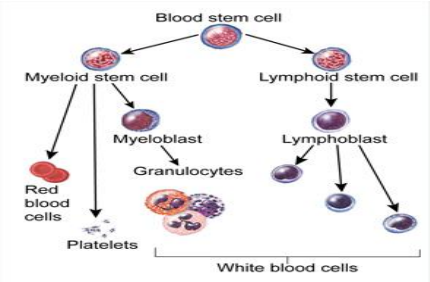

Fig1.All blood cells are produced by blood stem cells.

C. Leukemia Cells

In a person with leukemia, the bone marrow produces abnormal white blood cells that are called leukemia cells and leukemic blast cells. The abnormal cells can't produce normal white blood cells [9].Leukemia cells divide to produce copies of them. The copies divide again and again, producing more and more leukemia cells. Unlike normal blood cells, leukemia cells don't die when they become old or damaged. Because they don't die, leukemia cells can build up and crowd out normal blood cells[2]. The low level of normal blood cells can make it harder for the body to get oxygen to the tissues, control bleeding, or fight infections. Also, leukemia cells can

spread to other organs, such as the lymph nodes, spleen, and brain.

D. Types of Leukemia

Lab tests help the doctor find out the type of leukemia that you have. For each type of leukemia, the treatment plan is different.

Acute and Chronic Leukemias

Leukemias are named for how quickly the disease develops and gets worse:

- **Acute:** Acute leukemia usually develops quickly. The number of leukemia cells increases rapidly, and these abnormal cells don't do the work of normal white blood cells. A bone marrow test may show a high level of leukemia cells and low levels of normal blood cells. People with acute leukemia may feel very tired, bruise easily, and get infections often.
- **Chronic:** Chronic leukemia usually develops slowly. The leukemia cells work almost as well as normal white blood cells. People may not feel sick at first, and the first sign of illness may be abnormal results on a routine blood test. For example, the blood test may show a high level of leukemia cells. If not treated, the leukemia cells may later crowd out normal blood cells.

Myeloid and Lymphoid Leukemias

Leukemias are also named for the type of white blood cell that is affected:

- **Myeloid:** Leukemia that starts in myeloid cells is called myeloid, myelogenous, or myeloblastic leukemia.
- **Lymphoid:** Leukemia that starts in lymphoid cells is called lymphoid, lymphoblastic, or lymphocytic leukemia. Lymphoid leukemia cells may collect in the lymph nodes, which become swollen.

Four Most Common Types of Leukemia

- **Acute myeloid leukemia** (AML) affects myeloid cells and grows quickly. Leukemic blast cells collect in the bone marrow and blood.

  About 15,000 Americans will be diagnosed with AML in 2013. Most (about 8,000) will be 65 or older, and about 870 children and teens will get this disease.
- **Acute lymphoblastic leukemia** (ALL) affects lymphoid cells and grows quickly. Leukemic blast cells usually collect in the bone marrow and blood [7].

  More than 6,000 Americans will be diagnosed with ALL in 2013. Most (more than 3,600) will be children and teens.
- **Chronic myeloid leukemia** (CML) affects myeloid cells and usually grows slowly at first. Blood tests show an increase in the number of white blood cells. There may be a small number of leukemic blast cells in the bone marrow.

  About 6,000 Americans will be diagnosed with CML in 2013. Almost half (about 2,900) will be 65 or older, and only about 170 children and teens will get this disease.
- **Chronic lymphocytic leukemia** (CLL) affects lymphoid cells and usually grows slowly. Blood tests show an increase in the number of white blood cells. The abnormal cells work almost as well as the normal WBC.
  About 16,000 Americans will be diagnosed with CLL in 2013. Most (about 10,700) will be 65 or older. This disease almost never affects children or teens. Other, less common types of leukemia will account for more than 6,000 new cases in 2013

E. Risk Factors for Leukemia

The causes of leukemia are not known. However, several factors have been identified which may increase your risk. These include:

- family history of leukemia
- smoking (AML)
- genetic disorders such as Down syndrome
- blood disorders - myelodysplastic syndromes are sometimes known as pre-leukemia
- prior treatment for cancer with chemotherapy or radiation
- exposure to high levels of radiation
- chemical exposures, such as to benzene

F. Symptoms of Leukemia

The symptoms of leukemia include:

- excessive sweating, especially at night
- fatigue and weakness that do not go away with rest[5]
- unintentional weight loss
- bone pain and tenderness

- painless, swollen lymph nodes (especially in the neck and armpits)
- enlargement of the liver or spleen
- red spots on the skin (petechiae)
- bleeding and bruising easily
- fever or chills
- frequent infections

Leukemia can also cause symptoms in organs that have been infiltrated or affected by the cancer cells [7]. For example, central nervous system involvement can cause:

- headaches
- nausea and vomiting
- confusion
- loss of muscle control
- seizures

Leukemia can also involve the lungs, gastrointestinal tract, heart, kidneys, and testes

1. PROBLEM STATEMENT

Blood cell segmentation and identification is a vital in the study of blood as a health indicator. A complete blood count is used to determine the state of a person's health based on the contents of the blood in particular the white blood cells and the red blood cells. The main problem arises when massive amounts of blood samples are required to be processed by the hematologist or Medical Laboratory Technicians. Leukemia is a term used for diseases in which abnormal cells divide without control and are able to invade other tissues. Leukemic cells can spread to other parts of the body through the blood and lymph systems. Cancer is not just one disease but many diseases. There are more than100 different types of cancer. Most cancers are named for the organ or type of cell in which they start for example, cancer that begins in the colon is called colon cancer; cancer that begins in melanocytes of the skin is called melanoma. All cancers begin in cells, the body's basic unit of life. The time and skill required for the task limits the speed and accuracy with which the blood sample can be processed.

This project aims to provide user-friendly software based on MATLAB allowing for quick user interaction with a simple tool for the segmentation and identification of abnormal cells from a provided image.

2. HISTORY OF CANCER CELL AND ITS TREATMENTS

The earliest known descriptions of cancer appear in seven papyri, discovered and deciphered late in the 19th century. They provided the first direct knowledge of Egyptian medical practice. Two of them, known as the "Edwin Smith" and "George Ebers" papyri, contain descriptions of cancer written around 1600 B.C., and are believed to date from sources as early as 2500 B.C.

Hippocrates described several kinds of cancer, referring to them with the Greek word *carcinos* (crab or crayfish), among others.[1] This name comes from the appearance of the cut surface of a solid malignant tumour, with "the veins stretched on all sides as the animal the crab has its feet, whence it derives its name".[2] Since it was against Greek tradition to open the body, Hippocrates only described and made drawings of outwardly visible tumors on the skin, nose, and breasts. Treatment was based on the humor theory of four bodily fluids (black and yellow bile, blood, and phlegm). According to the patient's humor, treatment consisted of diet, blood-letting, and/or laxatives. Through the centuries it was discovered that cancer could occur anywhere in the body, but humor-theory based treatment remained popular until the 19th century with the discovery of cells.

Radiation Therapy

Radiation therapy or radiotherapy, often abbreviated RT, RTx, or XRT, is therapy using ionizing radiation, generally as part of cancer treatment to control or kill malignant cells. Radiation therapy may be curative in a number of types of cancer if they are localized to one area of the body. It may also be used as part of adjuvant therapy, to prevent tumor recurrence after surgery to remove a primary malignant tumor (for example, early stages of breast cancer). Radiation therapy is synergistic with chemotherapy, and has been used before, during, and after chemotherapy in susceptible cancers. The subspecialty of oncology that focuses on radiotherapy is called radiation oncology.

Chemotherapy

Chemotherapy is a category of cancer treatment that uses chemical substances, especially one or more anti-cancer drugs (chemotherapeutic agents) that are given as part of a standardized chemotherapy regimen. Chemotherapy may be given with a curative intent, or it may aim to prolong life or to reduce symptoms (palliative chemotherapy). Along with hormonal therapy and targeted therapy, it is one of the major categories of medical oncology (pharmacotherapy for cancer). These modalities are often used in conjunction with other cancer treatments, such as radiation therapy, surgery, and/or hyperthermia therapy. Some chemotherapy drugs are also

used to treat other conditions, including AL amyloidosis, ankylosing spondylitis, multiple sclerosis, Crohn's disease, psoriasis, psoriatic arthritis, systemic lupus erythematosus, rheumatoid arthritis, and scleroderma.

DNA

The greatest advances of the 20th century took place with the advent of the discovery of DNA by James Watson and Francis Crick. This led to a greater understanding of the detailed mechanisms of cancer

Normal Blood Counts

Normal blood counts fall within a range established by testing healthy men and women of all ages. The cell counts are compared to those of healthy individuals of similar age and sex. [3] Nearly all lab reports include a "normal" range or high and low "values" to help you understand test results.

Normal Ranges of Blood Cell Counts for Healthy Adults and Children of White Cells per *(μL)* micro liter of Blood

Men	5,000 to 10,000
Women	4,500 to 11,000
Children[3]	5,000 to 10,000

MEDICAL IMAGING

In the clinical context, medical imaging is generally equated to radiology or "clinical imaging ‘‘and the medical practitioner responsible for interpreting (and sometimes acquiring) the image is a radiologist. Diagnostic radiography designates the technical aspects of medical imaging and in particular the acquisition of medical images. The radiographer or radiologic technologist is usually responsible for acquiring medical images of diagnostic quality, although some radiological interventions are performed by radiologists. While radiology is an evaluation of anatomy, nuclear medicine provides functional assessment.

III. METHODOLGY

I) Three main techniques are used to estimate the number of RBC and WBC in the blood smear image

Which are logical, morphological and Hough transform. Fig 1 shows the complete flow chart of proposed system. The images are used in data acquisition stage are taken from online medical library. The input image needs to be pre-processed for further analysis.

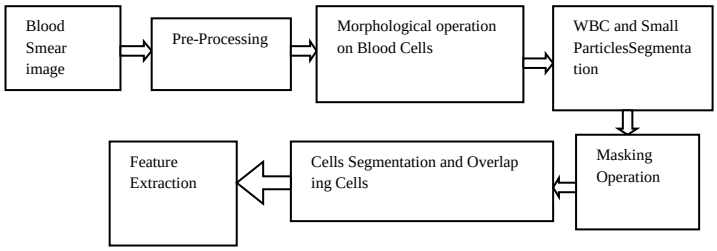

Fig2 Flow chart of Proposed system

II) The images from blood smear slides are captured by connecting high resolution digital camera to a microscope by adjusting microscope magnification to get good resolution. The pre-processing stage includes noise reduction and contrast enhancement of acquired image and is performed to prepare the image for the further stages. For processing the image it is converted into gray scale image, to avoid being influenced by dye color. A typical peripheral blood smear image consists of four components, namely background, erythrocytes, leukocytes, and thrombocytes.

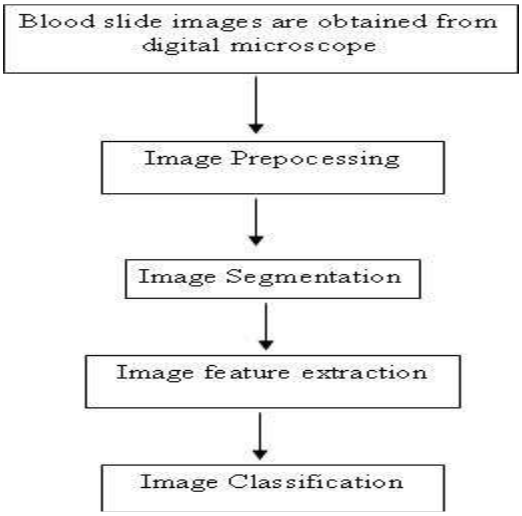

Fig3. Typical Steps In Process Of Automating Blood Recognition

Research methodology that will be used in this Research includes:

1 Image Acquisition

Blood image from slides will be obtained from Nearby hospital with effective magnification.

2 Preprocessing

During image acquisition and excessive staining, the images will be disturbed by noise. The noise may be due to illumination or shadows that make region of interest (ROI) appear as blurred image region. Background will be excluded since our ROI will be white blood cells. During this preprocess, image enhancement will be done as the contrast enhancement technique is capable to improve the medical image quality

3 Segmentation

Segmentation of white blood cell (WBC) and determine ROI that is nucleus for WBC only. This is because in leukemia cell images, the cytoplasm is scanty .So, focus will be on nucleus of WBC only. Determination the types of WBC should be done from the nucleus. Only lymphocytes and myelocytes should be considered and need to determine them whether they are blast cells or not. Others like neutrophils, basophils and eosinophils should be excluded. Once the blast cells are determined, then proceed to the next step. Sub images containing nucleus only will be considered. This is to reduce errors since there are similar color scales in WBCs with other blood particles.

4. Feature Extraction

The most important problem in generation of features of blood cells that characterize them in a way enabling the recognition of different blast types with the highest accuracy. The features to be used are for nucleus of lymphocytes and myelocytes:

· ***Geometrical Features*** – which includes area, radius, perimeter, symmetry, concavity, compactness, solidity, eccentricity, elongation, form factor will be obtained.
· ***Texture Features*** – which includes homogeneity, energy, correlation, entropy contrast, angular second momentum will be obtained.
· ***Color Features*** – the RGB color spaces will be transformed.
· ***Statistical Features*** – the mean value, variance, skewness.

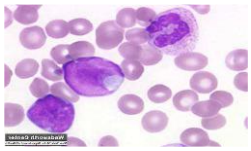

Fig 4. Acute Lymphocytes Leukemia

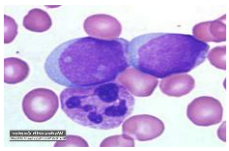

Fig 5. Chronic Myeloid Leukemia

IV. OUTLINE OF APPROACH

In this Approach we are using two different algorithms for the detection of the abnormal cells:

1. KNN(K-Nearest Neighbor Algorithm)
The K Nearest neighbor (kNN) decision rule has been a ubiquitous classification tool with good scalability. Past experience has shown that the optimal choice of K depends upon the data, making it laborious to tune the parameter for different applications.[4] The k-Nearest-Neighbors (kNN) is a non-parametric method of classification. It is simple but very effective in many cases[10]. Here also kNN has been utilized to classify blast cells from normal white blood cells.

HOUGH TRANSFORM

This is a feature extraction technique used in image analysis, computer vision, and digital image processing. The purpose of the technique is to find imperfect instances of objects within a certain class of shapes by a voting procedure. This voting procedure is carried out in a parameter space, from which object candidates are obtained as local maxima in a so-called accumulator space that is explicitly constructed by the algorithm for computing the Hough transform. The classical Hough transform was concerned with the identification of lines in the image, but later the Hough transform has been extended to identifying positions of arbitrary shapes, most commonly circles or ellipses The GHT is a modified version of the HT that not only searches for analytically defined shapes, but also arbitrary shapes (shapes that cannot be defined by an analytical equation). This method uses the principle of template matching, which relies on detecting smaller elements matching a template image The CHT set the radius to a constant value or provide the user with the option of setting (maximum and minimum) prior to running the application. For each edge point, a circle is drawn with that point as origin and radius. The CHT depends on a pre-define value of the circles radius

V. RESULTS AND DISCUSSION

As per both algorithms used in this research work, the outputs for an abnormal cells is given below,

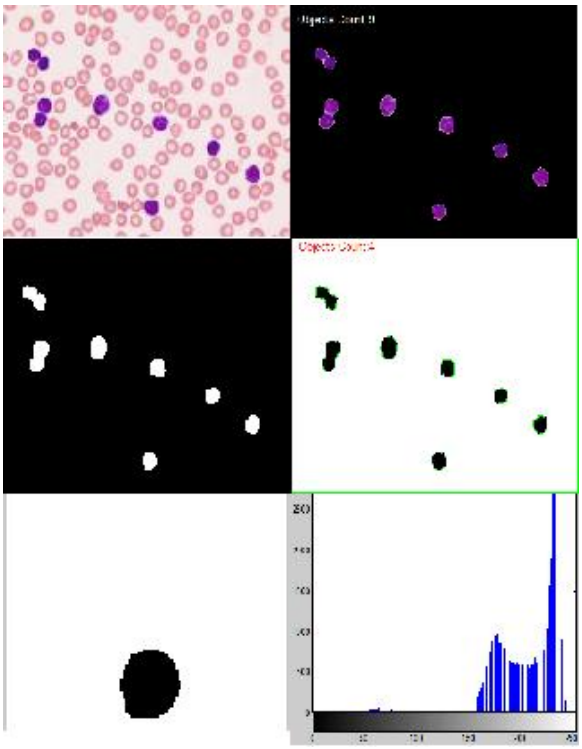

Fig 6: (i) input image, (ii) clustering and counting of abnormal cells by k-NN algorithm, (iii) Clustering using Hough Transform, (iv) objects counting by Hough Transform, (v) single cell extraction using MATLAB functions and (vi) Histogram plot for the input image.

The images in Figure 6 are output images for our research work, these images show the k-NN and Hough Transform algorithms implementation on input image shown in Fig 6 (i).

Evaluation Parameters

Area is calculated by using MATLAB commands; area is calculated by calculating Pixels in the black region.
Radius is calculated by using mathematical formula. Radius is calculated by square root of area divided by Pi.
Perimeter is calculated by mathematical formula i.e. 2 * Pi * Radius.
Time for execution of proposed algorithm is also calculated by using MATLAB function CLOCK.

TABLE OF READINGS

IMAGE NAME	AREA	RADIUS	PEREMETER
1.png	9191	54.08	339.84
2.png	6161	44.28	278.24
3.png	6161	44.28	278.24
4.png	2912	30.44	191.29

TABLE I: AREA, RADIUS and PARAMETERS.

IMAGE NAME	SHAPE	COLOR	TIME TAKEN
1.png	Irregular	Darkish	3.95
2.png	Sickle shape	Darkish	5.89
3.png	Tear drop	Light pink	2.53
4.png	Irregular	Light pink	2.33

TABLE II: SHAPE, COLOR and TIME TAKEN.
The images used for research work are shown below.

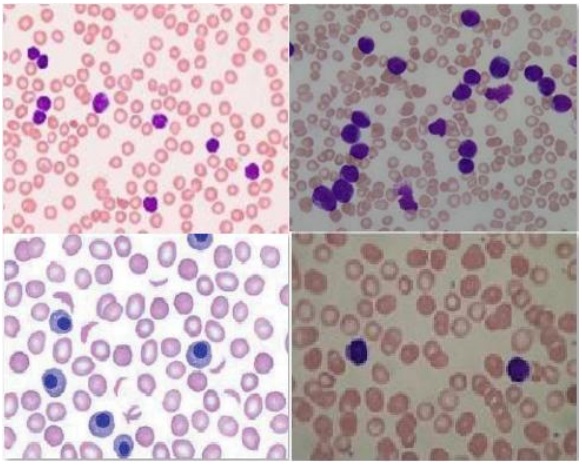

Fig 7: Input images used for research work.

VI. CONCLUSION AND FUTURE SCOPE

The paper has presented the image processing approach for the recognition and classification of the leukemia cells. The most important points of this approach are as: Different features of cells are then extracted from the labeled input image. From this feature, database is created and with the help of this database, set of test images and KNN classifier and HOUGH Transform, classification of various cells in an image has done. The main aim of this paper is blood slide image segmentation followed by feature extraction to detect leukemia cells. Shape features of nucleus such as area, perimeter, circularity, code implementation time etc. are considered for better accuracy of detection and extraction. Leukemia detection with proposed feature was classified using kNN classifier and Hough Transform giving overall accuracy of 93%. Furthermore the system should be robust to excessive staining and touching cells. Results obtained encourage future works which includes classification of lymphoblast

into various subtypes and counting, detecting other abnormal cells. In future someone can add some impressive GUI design to implement this code, or someone can use some other algorithm to implement the same design or calculate some other parameters.

REFERENCES

[1] RED BLOOD CELLS ESTIMATION USING HOUGH TRANSFORM TECHNIQUE in International Journal (SIPIJ) Vol.3, No.2, and April 2012.

[2] BLOOD CELLS EXTRACTION USING COLOR BASED SEGMENTATION TECHNIQUE International Journal of Life Sciences And Biotechnologies And Pharma Research, ISSN 2250-3137Vol. 2, No. 2, April 2013.

[3] IMAGE PROCESSING BASED APPROACH TO CANCER CELL PREDICTION IN BLOODSAMPLES. International Journal of Technology and Engineering Sciences Vol.1 (1), ISSN: 2320-8007.

[4] Classification of RBC and WBC in Peripheral Blood Smear Using KNN, PARIPEX - INDIANJOURNAL OF RESEARCH, Volume: 2 Issue: 1 | January 2013 ISSN - 2250-1991

.

[5] R.Gonzalez, R. Woods & Steven L, 'Digital image processing using MATLAB.

[6] White blood cell segmentation using morphological operators and scale-space analysis.

[7] National Cancer institute US http://www.cancer.gov

[8] W. Gonzalez, "Digital Image Processing".

[9] Medicine Health – http://www.medicinehealth.com/leukemia/article.htm

[10] **White** Blood Cells Segmentation and Classification to Detect Acute Leukemia.