IMAGE PROCESSING TECHNIQUES TO IDENTIFY RED BLOOD CELLS

Miss. Priyanka L. Khambayat[1], Dr. Dinesh D. Patil[2], Prof. Yogesh S. Patil[3]

[1]M.Tech, Computer Science & Engineering, Shri Sant Gadge Baba COET, Bhusawal, Maharashtra, India
[2]Head of Department and Associate Professor, Computer Science & Engineering, Shri Sant Gadge Baba COET, Bhusawal, Maharashtra, India
[3]Assistant Professor, Computer Science & Engineering, Shri Sant Gadge Baba COET, Bhusawal, Maharashtra, India

---***---

Abstract - *Sickle Cell Anemia is a blood disorder which results from the abnormalities in red blood cells of humans and shortens the life expectancy to 42 and 48 years for males and females. It also causes pain, jaundice, shortness of breath, etc. Sickle Cell Anemia is characterized by the presence of abnormal cells like sickle cell, ovalocyte, antisopoikilocyte. Sickle cell disease usually presenting in childhood, occurs more commonly in people from parts of tropical and subtropical regions where malaria is or was very common. A healthy RBC is usually round in shape, but sometime it changes its shape to form a sickle cell structure; this is called as sickling of RBC. Majority of the sickle cells whose shape is like crescent moon found are due to low hemoglobin content. An image processing algorithm is to automate the diagnosis of sickle-cells present in thin blood smears is developed. Images are acquired employing a charge coupled device camera connected to a light-weight microscope. Clustering based segmentation techniques are wont to identify erythrocytes (red blood cells) and Sickle-cells present on microscopic slides. Image features supported texture, color and therefore the geometry of the cells are generated, also as features that make use of a priori knowledge of the classification problem and mimic features employed by human technicians. The red blood corpuscle smears were obtained from IG Hospital, Rourkela. Image processing technique based identification of sickle-cells in anemic patient are going to be very helpful for automatic, sleek and effective diagnosis of the disease.*

Key Words: Anaemia, Sickle cell anaemia, Image Segmentation, Grey-Level Co-Occurrence Matrices (GLCM), K-means clustering

1. INTRODUCTION

Automated Sickle Cell Anemia Detector is a Mat-lab application that is diagnosing 'Sickle Cell Anemia' (drepanocytosis) from microscopic images of the patient's blood smear. It uses advanced image processing techniques to analyze the shape of 'Red Blood Cells' which are primary indicators of the disease. Our ultimate goal however, lies in automation of the pathological examinations, thereby expediting the process of differential diagnosis.

A red blood cell in normal physiological condition is circular in front view and bi-concave in side view. Sickle cell anemia is a hereditary blood disorder which primarily presents itself with high propensity for red blood cells to assume a cresentic or sickle-like shape. When the patient provides microscopic image of his/her blood sample to Automated Sickle Cell Anemia Detector, it uses edge detection algorithms to scan for the presence of abnormally shaped red blood cells in it. The Automated Sickle Cell Anemia Detector then proceeds to compare the ratio of normal RBC count to sickle shaped RBC count. A decision considering a threshold is then made to arrive at the conclusion to whether the patient is anemic or not. The application also provides a detailed report of its result for further diagnostic purposes if required. It also provides appropriate recommendations based on it.

1.1 Anaemia

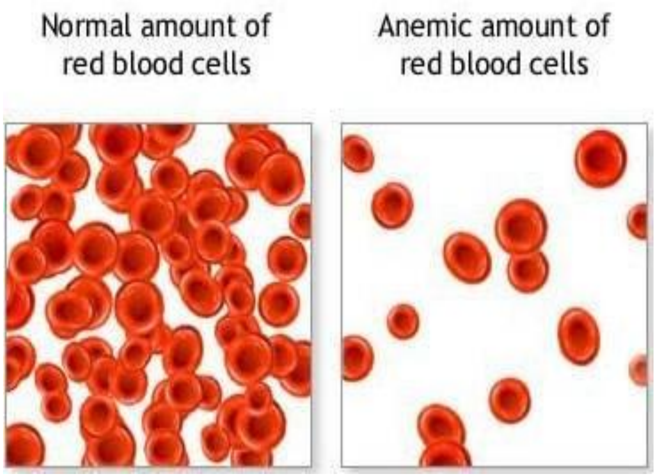

Fig -1: (a) Normal Blood Cell Image (b) Anaemic Blood Cell

The cellular part of blood molecule contains several different cell types in blood cells. One of the numerous cell and most important types are red blood cells. The other cell types are the white blood cells and platelets. Anemia is the most common disorder of the blood called Anemia. "Anemia", is the name is derivative from the ancient Greek word anemia, which means "Lack of Blood". It is possible because of reduction in Red Blood Cells (RBC) or resulting in lesser than normal quantity of hemoglobin in the human's blood. However, it can also include decreased oxygen- binding ability of each hemoglobin molecule due to deformity or lack of blood in numerical development. Anemia is actually a sign of a disease process rather than being a disease itself. It can be either classified as acute or chronic. In chronic anemia, symptoms typically begin slowly and progress gradually; whereas in acute anemia, symptoms can be abrupt and more distressing. Among many factors, nutritional (like vitamins and mineral deficiencies) and non- nutritional (like infection and hemoglobinopathies), that contribute to the onset of anemia; Iron Insuffiency and malaria plays a significant role. For men, anemia is typically defined as hemoglobin level of less than 13.5 g/dl and in women as hemoglobin of less than 12.0 g/dl.

1.2 Sickle-Cell

Sickle-cell disease (SCD), or sickle-cell anaemia (SCA) or drepanocytosis, is an autosomal recessive heritable blood cells in human body disorder with over dominance, categorized by red blood cells that assume an abnormal, stiff, sickle shape. The sickling occurs because of a mutation in the haemoglobin gene. It follows when a person inherits two abnormal genes (one from each parent) that cause their RBCs to change shape, which is similar to a crescent moon as shown in figure 2.

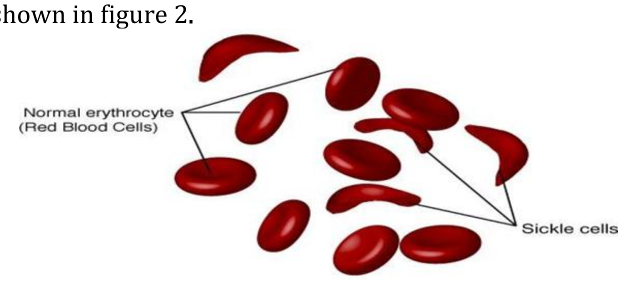

Fig -2: Sickle Cells and Normal Red Blood Cells

2. LITERATURE SURVEY

Sickle cell Anaemia is a well understood medical condition. There is no widely available standard cure. However there are certain treatments to prevent the symptoms and thus avoid complications resulting from it. These treatments require early detection as mentioned above. However there isn't any automated test for sickle cell Anaemia. This presumably stems from insufficiency of literature or studies involving automation of diagnostic process. Correspondingly, literature involving automated detection of sickle celled disease remains considerably scarce.

One of such recherché studies published on Science Direct is 'Detection of Abnormal Findings in Human RBC in Diagnosing Sickle Cell Anaemia Using Image Processing' authored by Pranati Rakshita and Kriti Bhowmikb [3]. This study treads along parallel ideologies as our project. The proposed methodology in it involves pre-processing, edge detection and region selection.

Pre-processing requires initially converting the blood smear image into a binary form. Then an adaptive filtering method is used to eliminate unwanted noise present in it. The filter of choice used in this paper was Weiner's filer. The next step is edge detection which demarcates boundaries for the red blood cells. The study approves the use of any of the following edge detectors: Sobel Operator, Robert's Operator, Canny Operator, Log Operator, Zero-cross Operator and Prewitt Operator [4]. The third step is region selection, wherein we measure properties of connection image components which satisfy certain predefined conditions. It is used to compute the shape measurements like the centroid, area, bounding box convex Hull, convex Area, perimeter etc. An extensive use of properties area and perimeter is done. Area is a scalar value which represents the actual number of pixels in the region and perimeter is used to calculate the distance around the boundary of the region. In MATLAB, these measurements can be computed using the inbuilt function 'Region Props'.

Using the measurements obtained, the study proposes a metric to determine the circularity of objects (which are the RBCs) in the image. This metric is defined as: (4*pi*area)/perimeter^2 [1]. The metric ranges from values of 0 to 1; 1 for a perfectly 2D circle, 0.785 for 2D square and so on goes decreasing as the shape resembles less of a circle. So a typical RBC would have a metric higher than 0.82 whereas a sickle shaped RBC would have a much lower metric of about 0.4 – 0.5. This we obtain a clear distinction between different shapes of RBCs present.

3. METHODOLOGY

3.1 Clustering Based Segmentation

1 Acquisition

Before examine the structure of RBCs, the images can be recorded with the help of glass slides and images get captured using microscopes. Images are acquired from IG Hospital, Rourkela using a 3.34 megapixel Nikon Coolpix 995 digital camera (Nikon Corporation, Tokyo, Japan). The camera, using full 4· optical zoom, is connected to a light microscope with 1,000· magnification. The blood smear image slides are examined under oil immersion. Images are captured in the JPEG format at the maximum resolution of the camera, 2,048·1,536 pixels.

2 Pre-Processing

The purpose of the pre-processing stage is to remove unwanted effects such as noise from the image, and transform or adjust the image as necessary for further processing. The resolution of the image is reduced by a factor of four to 512·384 to speed up performance of the system.

Also, the test images will be subjected to selective median filtering and un-sharp masking to isolate noise which may have been accumulated during image acquisition and due to excessive staining.

3 Color Transformation

Typically an image can be represented with the help of three color components. Images generated by the digital microscope are usually in RGB color space which is visually difficult to segment. For better color based segmentation we map the RGB image to L*a*b (LAB) color space. The LAB space consists of luminosity layer L*, chromaticity layers a* and b*. Since all the color information is in the chromaticity layers, we use these two components for color based Red Blood Cell segmentation.

4 Segmentation

Segmentation refers to the process of partitioning a digital image into multiple segments (sets of pixels, also known as super pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyse. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics. The anaemia blood smear will be segmented by appropriate segmentation methods.

5. Classification

After the various kinds of features were generated from the previous step, now they are used for classifying the different kinds of red blood cells (RBC) present in the blood smear image. After using proper grouping technique, we find that the form factor values extracted from the geometrical features is of our great interest. For different values of form factor, cells are sorted out as normal cell, sickle-cell, ovalocyte and anisopoikilocyte. Their numbers and the total number of blood cells present in the smear are also calculated. For a normal blood cell, the value of Form Factor nearly equals one. Consequently, the cells in blood smear are categorized as follows:

• Normal Cell	Form Factor > 0.95
• Sickle-Cell	0.6< Form Factor < 0.8
• Ovalocyte/Anisopoikilocyte	Form Factor < 0.5

4. ALGORITHM

4.1 K-Means Clustering

The K-means clustering, or Hard C-means clustering, is an algorithm based on finding data clusters in a data set such that a cost function (or an objection function) of dissimilarity (or distance) measure is minimized.

K-means clustering is a partitioning method. The function k-means partitions data into k mutually exclusive clusters, and returns the index of the cluster to which it has assigned each observation. Also, k-means clustering operates on actual observations (rather than the larger set of dissimilarity measures), and creates a single level of clusters.

K-means treats each observation in the data as an object having a location in space. It finds a partition in which objects within each cluster are as close to each other as possible, and as far from objects in other clusters as possible. You can choose from five different distance measures, depending on the kind of data you are clustering.

Each cluster in the partition is defined by its member objects and by its centroid, or center. The centroid for each cluster is the point to which the sum of distances from all objects in that cluster is minimized. K-means computes cluster centroids differently for each distance measure, to minimize the sum with respect to the measure that you specify. K-means uses an iterative algorithm that minimizes the sum of distances from each object to its cluster centroid, over all clusters. This algorithm moves objects between clusters until the sum cannot be decreased further. The result is a set of clusters that are as compact and well-separated as possible. You can control the details of the minimization using several optional input parameters to k-means, including ones for the initial values of the cluster centroids, and for the maximum number of iterations.

➢ **K-Means Clustering Algorithm**

1. Place K points into the space represented by the objects that are being clustered. These points represent initial group centroids.
2. Assign each object to the group that has the closest centroid.
3. When all objects have been assigned, recalculate the positions of the K centroids.
4. Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated.

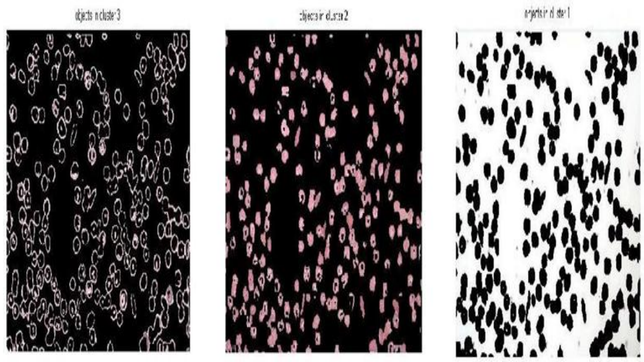

Fig -3: Objects in cluster after K-Means Clustering based segmentation

4.2 Fuzzy C-Means Clustering

Fuzzy c-means (FCM) is one of the commonly used methods for image segmentation and its success is mainly due to the introduction of fuzziness for the belongingness of each image pixels. Compared with crisp or hard segmentation methods, FCM is able to retain more information from the original image. However, one disadvantage of FCM is its sensitivity to noise and other imaging artifacts. The reason may perhaps stems from the non-unimodal property of its membership functions and the use of the squared Euclidean distance.

Fuzzy c-means (FCM) is a data clustering technique in which a dataset is grouped into n clusters with every data point in the dataset belonging to every cluster to a certain degree. For example, a certain data point that lies close to the center of a cluster will have a high degree of belonging or membership to that cluster and another data point that lies far away from the center of a cluster will have a low degree of belonging or membership to that cluster.

- **Fuzzy C-Means Clustering Algorithm:**

Assumptions: Image transformed into feature space, number of clusters c, stop condition ε, fuzziness parameter m.

Step 1 Cluster image in feature space, with next conditions: number of clusters is c, fuzziness index is m and stop condition is ε.

Step 2 Repeat for each pixel aij of image I.

Step 2.1 Find out, into which cluster Cl belongs pixel aij at most.

Step 2.2 Find out, whether in the closest surroundings of pixel aij exists segment Rk, which points belong to same cluster C.

Step 2.3 If such segment Rk exists, than pixel aij add to segment Rk, else create new segment Rn and add pixel aij to new segment Rn.

Step 3 Merge all segments, which belong to one cluster and are neighbors

Step 4 Arrange borders of all segments.

5. SYSTEM ARCHITECTURE

Automated Sickle Cell Anaemia Detector provides a solution that is not just fast and inexpensive, but also can be easily incorporated into routine blood-work performed at birth. It also expedites the diagnosis of sickle cell anaemia in adult patients, by circumventing the lengthy manual blood examinations. Additionally, it resolves the possibility of human-induced error that may occur during manual testing. The system model includes two parts: the system server and user. In this module the system server is pathologist which takes high quality microscopic image of blood smear of the patient to check abnormal sickle cells, which uses image processing techniques as image acquisition, image segmentation, feature extraction and feature classification. after that microscopic image will converted to binary image then identify large RBCs individually to calculate area and perimeter of each RBC then calculate the metric and classify all sickle cells and displays the final output as total number of objects detected, number of valid cells, number of normal cells, number of abnormal cells, threshold and result declared as you are safe or unsafe.

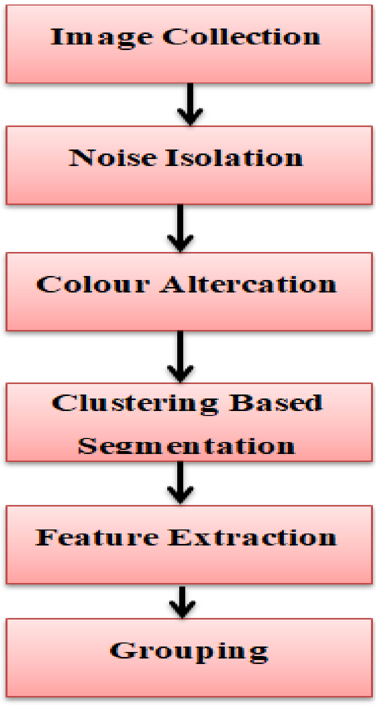

Fig -4: System Overview

6. RESULT

The following are the results obtained after providing an input image with blood smear containing sickle cells image.

6.1 Result of histogram stretching

Histogram stretching (or normalization), is the process of improving or enhancing the contrast of an image. Contrast is the difference between maximum and minimum pixel intensity. This is performed in the software to obtain a clear distinction between a cell and its background.

- **Input:** A grey scale image with poor contrast.
- **Expected Output:** Improving contrast of the image having poor picture quality
- **Observed Output:** Better dynamic range and noticeable enhancement in contrast image which is more clear and easy to analyse the final output that is whether patient is safe or not..

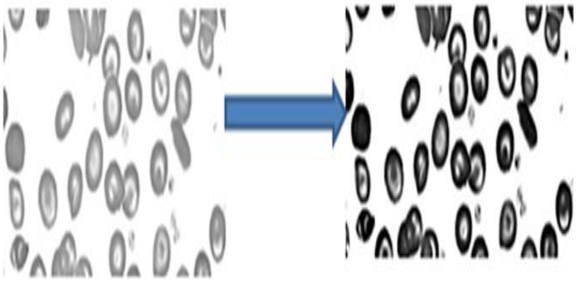

Fig -4: histogram stretching

Following snapshots are of the proposed system which shows the execution of the system.

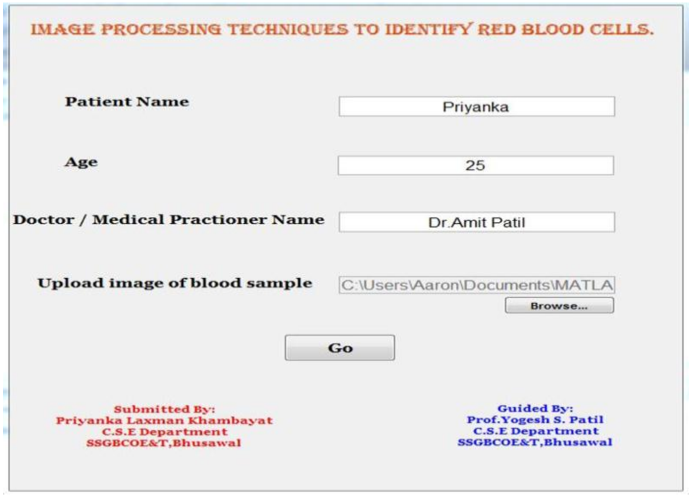

Fig -5: Filling the patient details and uploading image

In the above figure 5 shows the fields as Patient name, age, Doctor or practitioner name with upload image of blood sample. This proposed system is able to identify red blood cells using image of the sample blood cells. After click on Go button it process the image while processing image it converts the real image into binary image form contain grey scale image. After stretching the image shows the red cells clear and binary image contain red and black dots where black dots are the red cells.

After process the image at the end it shows the final result as shown in the figure 6. In result I, Observed that it gives all the patient details with shows total number of valid cells, total number of normal cells. Finally shows the result patient is safe or not by analyse the uploading image by comparing sample pattern data set in the system.

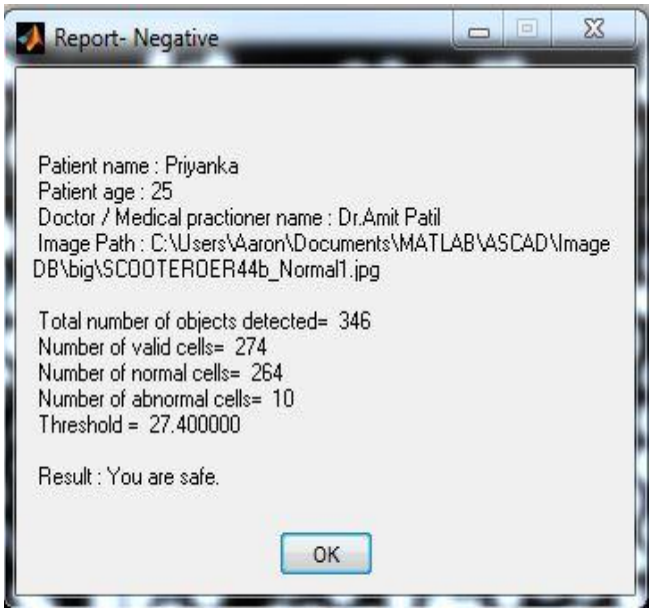

Fig -6: Showing result with details

7. CONCLUSION

India is a country with a population of over 1.27 billion people. It is estimated that out of this, 200 million suffer from some kind of disease at any given point of time. On top of the world second largest sick population, a poor health-care funding and lack of R&D in medicine paints a deplorable picture of the prevalent health-care standard in our country. In such a scenario, our solution to sickle-celled disease is what we believe, an initiative in the right direction. By expediting the diagnostic processes, we not only seek to improve the health individually but also curtail unnecessary industry-wide expenditure involved in long & manual laboratorial work. Furthermore, Automated Sickle Cell Anaemia Detector, if successful as a consumer product, has the potential to open up a gateway to private research and development investment into automation in health-care sector. Such solutions that target the root causes are indispensable in improving the health-care standard in our country.

8. FUTURE SCOPE

Using the image processing and segmentation technique also the following sub-imaging technique, I can obtain the images of particular affected RBCs, i.e. Sickle-cells, Anisopoikilocytes and Ovalocytes and further apply feature extraction process to determine the characteristics of affected RBCs and thus make an artificial neural network to automatically diagnose sickle-cells disease affected person. This proposed system is useful to detect red blood cells automatically without specialist involvement.

REFERENCES

[1] D.J Weatherall, J.B Clegg, "Inherited haemoglobin disorders: an increasing global health problem" Bull. World Health Organ. 79: 704–12, 2001

[2] N Awasthy, K. C. Aggarwal, P. C. Goyal,M.S Prasad, M. Sharma, "Sickle cell disease: Experience of a tertiary care centre in a nonendemic area" Ann Trop Med Public Health, 2008.

[3] P. Rakshita, K. Bhowmikb, "Detection of Abnormal Findings in Human RBC in Diagnosing Sickle Cell Anaemia Using Image Processing", International Conference on Computational Intelligence: Modeling, Techniques and Applications, Elsevier, 2013

[4] N. S. Aruna, S. Hariharan, "Edge Detection of Sickle Cells in Red Blood Cells", International Journal of Computer Science and Information Technologies, Vol. 5 (3) 4140-4144, 2014.

[5] S. S. Barpanda, "Use of Image Processing Techniques to Automatically Diagnose Sickle-Cell Anaemia Present in Red Blood Cells Smear", NIIT Rourkela, 2013.

[6] NCMH Background Papers—Burden of Disease in India (New Delhi, India), September 2005, Ministry of Health & Family Welfare, Nirman Bhawan, Maulana Azad Road, New Delhi 110011, India.