Analysis of White Blood Cells for Malaria Detection

Hemalatha K L, Karthik R, Nandan Sateesh Hegde

Dept.of ISE, SKIT, Bangalore,

hema.sjcit@gmail.com, karthikrajendra4@gmail.com, nandan.hegde678@gmail.com

ABSTRACT— Malaria is a diseases caused by bacteria, viruses or parasites transmitted by mosquitoes. It is a serious global health problem. The current way of malaria parasite diagnosis is microscopic observation of the blood slide by a trained and experienced lab technician which takes more time and expensive due to the complex contents of the blood smear and the expertise requirement. This paper proposes a combination of color and morphological based segmentation techniques for blood cells segmentation on microscopic images. K-means clustering is implemented for color based clustering and primary segmentation of the image. The clustering is cascaded into two layers for cell recognition and background elimination. The proposed method overcomes the problem of previous work which is over segmentation and efficiently segments the WBC cells and keeps cytoplasm unaffected.

Keywords-blood cells segmentation; k-means clustering;

I. INTRODUCTION

Malaria is one of the predominant tropical diseases in the world causing wide spread sufferings and deaths in the developing countries[4]. In 2010, about 3.3 billion people which are half of the world populations are at risk of malaria. In addition, this disease has caused the death of an estimation of 655,000 people in 2010, with 86% of the victims are children under five years of age [1]. The proportion of counts of different types of white blood cells in the bone marrow, called differential counts, provides invaluable information to doctors for diagnosis. Due to the tedious nature of the differential white blood cell counting process, an automatic system is preferable [2].

The need for automation of leukemia detection arises since current methods involve manual examination of the blood smear as the first step toward diagnosis. This is time-consuming, and its accuracy depends on the operator's ability [3]. The clinical behavior of the disease can be predicted using this classification and accordingly treatment should be given to the patient. In leukemia disease, large numbers of abnormal white blood cells are produced by bone marrow due to unknown cause. In pathology manual detection of leukemia is done which is time consuming as well as costly due to high

cost pathology instruments. Hence automatic technique is adopted for fast and accurate results [5].
This paper proposes a combination of color and morphological based segmentation methods for blood cells segmentation on microscopic bone marrow images. The main goal of this paper is to reduce over segmentation phenomenon by using k-means algorithm. Kmeans clustering is the color based technique performed as a primary segmentation stage to cluster the image pixels according to the similarity of color components and spatial coordinates.

One of the disadvantages of diagnosis using manual microscopy methods is that it requires extensive human

intervention during the diagnostic process which can often lead to late and sometimes erroneous diagnosis. The microscopist requires extensive training to gain expertise in the diagnosis, and because of the sheer volume of the samples that need to be analysed, the method is not consistent and is dependent upon blood smear and stain quality, microscope quality and the expertise of the microscopist. [7]

Some of the problems of manual microscopy can be overcome by exploring computer based, specifically image-based, diagnostic methods. The aim of this study is to outline a semi-automatic diagnosis method based on image processing and one that provides a reliable and consistent solution. The literature contains descriptions and details of several computer vision or image-based algorithms[8].

As malaria is a health care issue and is causing a serious health problem, the detection in the early stage is very important.

II. METHODOLOGY

The proposed procedures to develop a model for the Analysis of white blood cells for malaria detection overall working principle is depicted in Fig. 1. The steps are summarized as follows:

Step 1: Read the malaria slide images as the input.

Step 2: (Red Green and Blue) RGB to Grayscale conversion.

Step 3: Image filtering and thresholding.

Step 4: Erosion and Dilation

Step 5: Apply K-means for image segmentation.

Step 6: Labeling and counting.

STEP 1: Read the malaria slide images as the input:

In this step, the stained blood slide is captured and is given as an input to the System. The images are usually collected from the labs and the Hospitals for detection of malaria.

Step 2: RGB to Grayscale conversion:

Once the image is read into the system, it is converted to a Grayscale image. The contrast of the gray scale image is enhanced using local histogram equalization to enhance the visibility of the parasites and RBC.

Step 3: Image filtering and thresholding:

Filtering is done to reduce noise within an image or to produce a less pixilated image. Most filtering methods are based on low pass filters. There might be some unwanted regions or noises that are still encountered in the image. Thus, median filter is used as a noise removal in order to obtain a noise-free image

Thresholding is the simplest method of image segmentation. From a grayscale image, thresholding can be used to create binary images. The input to a thresholding operation is typically a grayscale or color image. In the simplest implementation, the output is a binary image representing the segmentation. Black pixels correspond to background and white pixels correspond to foreground (or *vice versa*). In simple implementations, the segmentation is determined by a single parameter known as the *intensity threshold.* In a single pass, each pixel in the image is compared with this threshold. If the pixel's intensity is higher than the threshold, the pixel is set to, say, white in the output. If it is less than the threshold, it is set to black.

Step 4: Erosion and Dilation:

The most basic morphological operations are dilation and erosion. Dilation adds pixels to the boundaries of objects in an image, while erosion removes pixels on object boundaries. The number of pixels added or removed from the objects in an image depends on the size and shape of the *structuring element* used to process the image. In the morphological dilation and erosion operations, the state of any given pixel in the output image is determined by applying a rule to the corresponding pixel and its neighbors in the input image. The rule used to process the pixels defines the operation as dilation or erosion.

Dilation is one of the two basic operators in the area of mathematical morphology, the other being erosion. It is typically applied to binary images, but there are versions that work on grayscale images. The basic effect of the operator on a binary image is to gradually enlarge the boundaries of regions of foreground pixels. Thus areas of foreground pixels grow in size while holes within those regions become smaller.

Erosion is the other operator of mathematical morphology, the basic effect of the operator on a binary image is to erode away the boundaries of regions of foreground pixels (*i.e.* white pixels, typically). Thus areas of foreground pixels shrink in size, and holes within those areas become larger.

Step 5: Apply K-means for image segmentation:

After transforming the RGB into grayscale, the next and important step in image segmentation is to extract the meaningful region from malaria image.

Clustering is a method to divide a set of data into a specific number of groups. It's one of the popular method is k-means clustering. In k-means clustering, it partitions a collection of data into a k number group of data [6]. It classifies a given set of data into k number of disjoint cluster. K-means algorithm consists of two separate phases. In the first phase it calculates the k centroid and in the second phase it takes each point to the cluster which has nearest centroid from the respective data point. There are different methods to define the distance of the nearest centroid and one of the most used methods is Euclidean distance.

Once the grouping is done it recalculate the new centroid of each cluster and based on that centroid, a new Euclidean distance is calculated between each center and each data point and assigns the points in the cluster which have minimum Euclidean distance. Each cluster in the partition is defined by its member objects and by its centroid. The centroid for each cluster is the point to which the sum of distances from all the objects in that cluster is minimized. So K-means is an iterative algorithm in which it minimizes the sum of distances from each object to its cluster centroid, over all clusters.

K-Means algorithm:

K-means clustering[9] is an iterative, data-partitioning algorithm that assigns n observations to exactly one of k clusters defined by centroid, where k is chosen before the algorithm starts.
The algorithm proceeds as follows:

1. Choose k initial cluster centers (centroid)
2. Compute point-to-cluster-centroid distances of all observations to each centroid.
3. There are two ways to proceed :
   - **Batch update** — Assign each observation to the cluster with the closest centroid.
   - **Online update** — Individually assign observations to a different centroid if the reassignment decreases the sum of the within-cluster, sum-of-squares point-to-cluster-centroid distances.
4. Compute the average of the observations in each cluster to obtain k new centroid locations.
5. Repeat steps 2 through 4 until cluster assignments do not change, or the maximum number of iterations is reached.

Step 6: Labeling and counting:

Based on the K-means clustering algorithm, the cells that are affected by malaria can be detected. We can label the cells that are affected by malaria and also the total number of cells that have been tested. The number of blood images affected by malaria. We can label the cells affected by malaria and a count of cells affected.
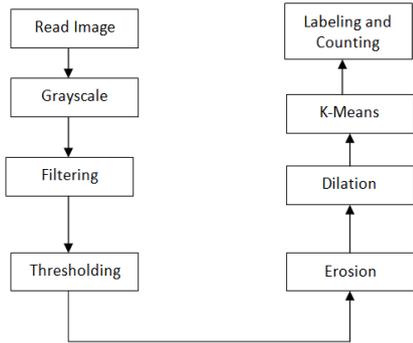

Figure 1
Block diagram of the proposed methodology

I. RESULT

The proposed technique has been applied on few peripheral blood smear images obtained from the public dataset. A microscopic blood image of size 700×468 and jpeg (figure 2) format is considered for evaluation. The image is first converted into a grayscale image as shown in figure 3. Once the image is converted into the Grayscale image, the image is clustered. The clustered image is shown in figure 4. The image is then segmented (figure 5). The result also contains the cells affected by Malaria parasite. Figure 6 shows the screen shot of the final result, the total number of cells tested were 23 and the affected cells were 4
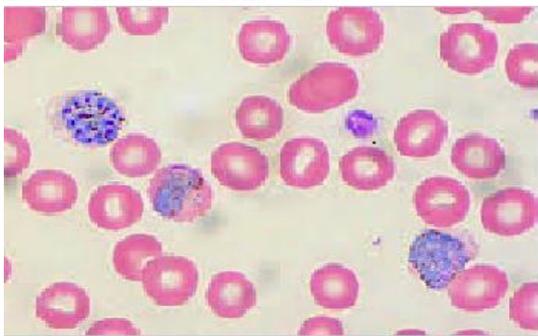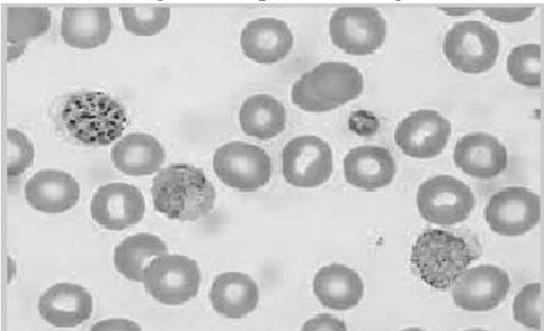

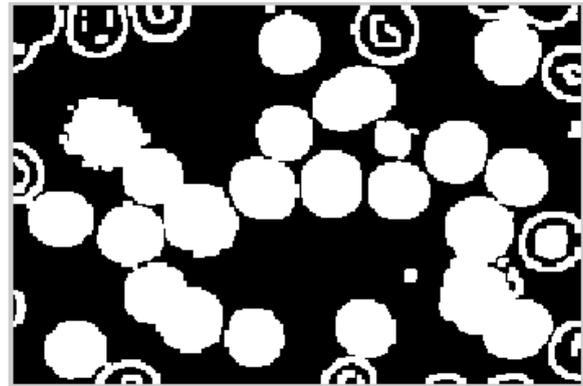Figure 2: Input Cell Image

Figure 3: Grayscale Image
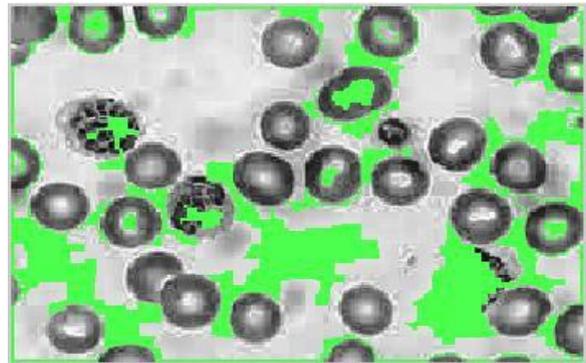

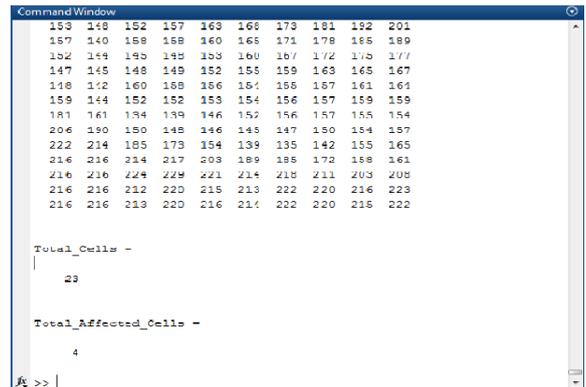Figure 4: Clustered Cell Image

Figure 5: Segmented Cell Image

Figure 6: Screen shot of the final output

II. CONCLUSION

The image segmentation is done using k-clustering algorithm, by initially creating the k centroid. We have made the erosion and Dilation which improves the quality of original image. The filtering of the image is done using the median filter. Hence, we can conclude that the proposed clustering algorithm has better accuracy and is more robust in nature.

III. FUTURE WORK

The methodology can be improved to obtain more efficient and quicker results. The k-means algorithm can be modified with the combination of other algorithms to obtain better system. Also, we can improve the quality of the output

image more by using the morphological operation and get better performance measurement. We can also implement different clustering method using subtractive clustering algorithm and lastly we can implement and analyze in different areas of image segmentation. The quality of the image can be improved and the system can be made more accurate with better efficiency.

IV. REFERENCES

[1] WHO, World Malaria Report 2011, World Health Organization, 2011.

[2] N. Theera-Umpon, S. Dhompongsa (2007). Morphological granulometric features of nucleus in automatic bone marrow white blood cell classification, in Information Technology in Biomedicine, IEEE Transactions on, 2007.11(3): pp 353-359.

[3] S. Agaian, M. Madhukar, A.T. Chronopoulos (2014). Automated screening system for acute myelogenous leukemia detection in blood microscopic images, in IEEE Systems Journal.

[4] The Malaria Control Programme, World Malaria report, World Health Organization, 2008

[5] H.N. Lim, M. Mashor, R. Hassan (2012). White blood cell segmentation for acute leukemia bone marrow images, in Journal of Medical Imaging and Health Informatics Volume 2, Issue-3, May–June 2013.

[6] R. Yager and D. Filev, Generation of Fuzzy Rules in Mountain Clustering, In Journal of Intelligent and Fuzzy System, vol. 2(3), pp. 209–219, (1992).

[7] Yashasvi Purwar, Sirish L Shah, Gwen Clarke, Areej Almugairi, and Atis Muehlenbachs (2013). Automated and unsupervised detection of malarial parasites in microscopic images, in Journal of Malar J. Volume 10, Dec 2013.

[8] Rao KNRM, Dempster AG, Jarra B, Khan S. Automatic scanning of malaria infected blood slide images using mathematical morphology. Proc IEE Semin Med Appl of Signal Process, London, UK.2002.

[9] Lloyd, Stuart P. "Least Squares Quantization in PCM." IEEE Transactions on Information Theory. Vol. 28, 1982, pp. 129–137.