Malarial Red Blood Cell Counter

EE368 Project Proposal
Austin Zheng
austinz@stanford.edu

Project Description

Malaria, a blood-borne disease transmitted by mosquitoes, involves the infection of red blood cells in humans and other organisms by protists of the genus *Plasmodium*. In order to diagnose and characterize a malarial infection for medical or scientific purposes a blood sample is drawn, smeared onto a slide, and stained with a nucleus stain. Because mature red blood cells do not possess cell nuclei, the stain only strongly marks the malarial parasites. This technique can be used to determine what proportion of red blood cells in the sample have been infected by counting the number of infected and uninfected red blood cells.

A red blood cell is considered infected if at least one, but possibly multiple parasites can be detected within its interior. White blood cells and free-floating parasites are not considered. The current state of the art involves manual counting by a laboratory technician or other individual, who can distinguish staining artifacts from actual nuclei, white blood cells, and (depending on specific requirements) life cycle and species of malarial parasites [1]. Although manual counting is relatively inexpensive to implement (requiring only basic laboratory equipment), adequate sensitivity requires proper training and supervision of technicians. This poses problems for both medical care providers in impoverished regions of the world as well as laboratory settings which may benefit from automation of a tedious and time-consuming task [2]. Conceivably, automation of this task could both facilitate laboratory efficiency as well as provide an alternative diagnostic tool in conjunction with mobile phone based microscopy in developing countries [3].

This project will attempt to produce a practical MATLAB replacement for manual counting, and also provide a platform for further refinement and extension based on the evolving needs of the laboratory. Efficacy against parameters will be measured using false positive and false negative detection rate against a manually labeled set of ground truth images. Examples of future improvements may include outputting an annotated image to an interactive front-end to allow for human verification, using image data to distinguish between parasites according to lifecycle stage, and utilizing human training data to allow for improved recognition accuracy using machine learning techniques.

Technical Approach

Processing will be carried out using MATLAB. The input will consist of color image files captured from a microscope and digitized. An example image is presented in Figure 1

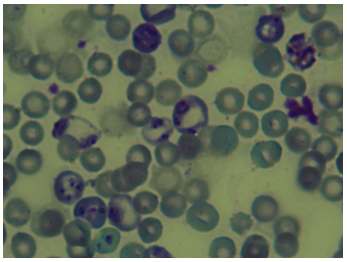

Figure 1: Blood smear

The first stage of image processing will involve background/foreground differentiation through the use of color. The background of each image (comprised mainly of the blood cells) is colored green, while potential areas of interest ('foreground') are colored a shade of purple which varies in intensity from moderately dark to extremely dark. Given the highly distinct contrast between background and foreground and the lack of other, irrelevant color information thresholding using HSV or RGB should be sufficient to separate out areas of interest.

The second stage of image processing will involve region labeling and characterization. MATLAB's `bwlabel` and `regionprops` functions provide a set of powerful tools for determining properties of regions, including area, convex hull, centroid, position within the image, eccentricity, and angular orientation. Region characterization in conjunction with color intensity information should allow for parasite nuclei to be identified and for erroneous signals (stain color on irrelevant regions due to imperfections in proessing) to be rejected. A well-designed framework for this stage should allow the system to be easily extended in the future to characterize detected parasites according to life-cycle stage based on their shape and size.

The third stage of image processing will involve determining the locations and positions of red blood cells. A reasonably clear separation between the edges of red blood cells and the background of the slide can be observed. Edge detection (using the Canny algorithm) in conjunction with the Hough transform as applied to circles should be sufficient to isolate regions of interest, and similar techniques have been used to perform generalized cell detection [4]. Circles whose radii are significantly larger or smaller than the average detected radius can be rejected as either free-floating parasites or white blood cells. The output of this stage will be data indicating which regions upon the image correspond to the interiors of individual red blood cells.

With data from these three stages it should then be possible to identify and count the number of red blood cells infected with malarial parasites, and to provide this information for use elsewhere within the experimental or diagnostic pipeline.

References

[1] D.C. Warhurst, J.E. Williams, *Laboratory diagnosis of malaria*, ACP Broadsheet No 148, 1996.

[2] P. Guerin, P. Olliaro, F. Nosten, P. Druilhe, R. Laxminarayan, F. Binka, W. Kilama, N. Ford, N. J. White, *Malaria: current status of control, diagnosis, treatment, and a proposed agenda for research and development*, The Lancet, Infectous Diseases vol 2, p. 566, 2002.

[3] Breslauer DN, Maamari RN, Switz NA, Lam WA, Fletcher DA (2009) Mobile Phone Based Clinical Microscopy for Global Health Applications. PLoS ONE 4(7): e6320. doi:10.1371/journal.pone.0006320

[4] T. Nattkemper, W. Schubert, T. Hermann, H. Ritter, *A hybrid system for cell detection in digital micrographs*, 2004.