Team NOMADD
Proposal (Fall 2014)

Daniel Cohen, Gregory Cooke, Phoebe Edalatpour, Garrett LeCroy, Kunal Mehan

How Might We…

"How might we diagnose malaria severity in the field with minimal expertise, improved accuracy, and reduced cost?"

Background

Malaria is a potent disease. Even with modern medicine and treatments, malaria continues to wreak havoc across the majority of tropical countries. Of the more than 1 million people die who each year from malaria; most are children under the age of 5 in sub-Saharan Africa [1]. Nearly half of the world's population lives in regions where malaria is endemic [2]. Although four main strains of malaria exist, *Plasmodium falciparum* is the most deadly and prevalent strain [3].

Malaria is caused by parasitic protozoa that infect humans through mosquito carriers. These organisms first reside in the liver of the infected human, then spread to the bloodstream roughly 14 days after infection. These parasites then invade the blood cells and reproduce asexually in a regular cycle whose period depends on the strain of malaria [1].

The threshold levels of malaria parasite density necessary to cause illness vary across populations; thus, accurate parasite concentration data is vital in determining the treatment plan for the patient. Patients from non-endemic regions will need anti-malarial medication at much lower parasite densities [4].

Early detection is the most valuable asset in the effort to stop malaria, because there are no proven methods to completely prevent malaria [1]. According to the NIH, the lack of early diagnosis is actually the leading cause of death from malaria [2]. In addition, the misdiagnosis of malaria has led to overuse of anti-malarial medicine, causing resistant strains to appear [5].

The current standard for malaria diagnosis is microscopy [1]. Microscopy allows technicians to actually see the malaria parasite in human blood samples. Current microscopy based methods of detection require expensive equipment and trained technicians. Microscopy does, however, allow for a definitive malaria diagnosis, and is also able to detect concentrations as low as 5 parasites per μL [3].

In the previous five years, researchers have proven that it is possible to use computer vision to reliability detect and classify malarial parasites [6]. However, this work has yet to be applied outside the lab. This is likely due to difficulties in implementing the algorithms with low cost imaging and computing hardware, a problem which is being increasingly mitigated by the recent glut of low cost, yet powerful hardware.

People living in remote, underdeveloped areas with poor access to immediate diagnostic services will be the demographic most affected by rapid, reliable diagnosis of malaria. Often, people in these areas are unable to detect symptoms until the disease has progressed to an advanced stage, at which point treatment is significantly more difficult . Because of this, each day, nearly 1,500 children die of malaria in underserved regions of Africa [7].

The economic impact of malaria is devastating to developing nations. Leighton and Foster [8] estimate that the total annual value of malaria-related production losses was 2 to 6 percent of GDP in Kenya and 1 to 5 percent in Nigeria, a significant and preventable burden.

A cheap, reliable system for malaria diagnosis has the potential to address many of the aforementioned problems with very little societal investment, garnering support from all stakeholders.

Project Goals

We aim to develop a vision analysis algorithm that can detect malaria parasites in blood smear samples and quantify the density of the parasites enabling earlier detection of malaria than antigen based tests. Early detection of malaria is crucial to providing effective care [1,2]. In the early stages of infection, malaria parasite densities in patients' blood samples, are low (200 parasites per μL), so microscopy is the most effective way to detect the malaria parasite [3]. Microscopy is a relatively expensive technique, since a trained technician is needed to analyze the blood smear samples. By utilizing already developed vision analysis algorithms as foundations, along with low cost optics, we can create a cheaper, faster method of detecting malaria in blood smear samples. Our proposed software could increase the ease of malaria detection, particularly in remote endemic regions. This algorithmic analysis will also help healthcare providers treat malaria more effectively due to improved accuracy in parasite density measurements, therefore decreasing the number of malaria related deaths.

External Advisors

There are a number of professors at Georgia Tech who conduct research pertinent to our project. We intend to reach out to them and ask if and how their expertise could aid our project process. We have specifically been looking for professors in BME, ECE, and CS to help us. In particular, Allen Tannenbaum, a joint BME/ECE professor that focuses on computer vision, biomedical imaging, control theory, and image processing would be a great resource if he is available to help us. His lab is the biomedical imaging lab. Additionally, Baowei Fei is a BME professor who specializes in biomedical imaging. He is a member of the Tech faculty, but his offices and labs are located at Emory.

We also have connections with advisors who do not work at Tech. We intend to get in contact with Karol Hatzilias, because his expertise and entrepreneurship experience would be invaluable to our project. We've also had brief correspondence with Eric Kamont, a solution specialist for Microsoft's Internet of Things. He has offered to try to provide help on the software development end of our project, if possible. Lastly, we have visited the CDC once already and spoken with malaria experts with whom we plan to remain in contact. Their knowledge of malaria and detection methods has been crucial in defining our problem statement.

Objectives

The objectives represent steps in an iterative process. As such, we will have very little, if any way to evaluate our progress until we have completed Objective III. At this point, we will have enough hard numerical evidence to decide what direction to pursue for the next iteration. This iterative process can be continued until the completion requirements outlined in Objective III are satisfied.

Objective I

Assess applicability of existing algorithms and techniques to malaria identification

Rationale

Human microscopy technicians become proficient at identifying malarial parasites through years of practice, gradually forming a mental model of what makes a given object a parasite. In order for a computer to accomplish the same task, it too must form a model of which objects are parasites, and which are not. This problem falls into a branch of Computer Science known as Machine Learning, and more specifically Classification. Computer Scientists have developed a multitude of algorithms to categorize objects into two or more classes, but some of them are better suited to specific applications than others. Without an understanding of existing techniques for this type of problem, it would be difficult if not impossible to implement a working technique for malaria classification.

Task I: Narrow the Field

We must narrow down the field to settle on an algorithm that will lend itself to malarial identification. A first pass can be made on intuition alone. A quick review of the literature will enable us to definitively exclude certain techniques because they don't work with the type of data we're using, or they don't allow us to supervise their learning. It may be difficult to formulate a definitive list, since this could easily consume a substantial amount of time, and many unknowns will remain until a chosen algorithm is actually implemented.

Measuring Completion

It is extremely difficult to further narrow our choices without some form of experimentation. Since we have the resources to construct a fully functional classifier in a short amount of time at almost zero cost, we will try each of the techniques that have a chance of working, intelligently transitioning between them as we learn more about the problem space. The final evaluation of our success will be measured by the performance of the technique during the testing phase. By quickly going to a working implementation, we can make our mistakes quickly, and rapidly converge on the optimal solution.

Objective II

Implement algorithm(s)

Rationale

As mentioned in the previous section, given the easily available resources to test algorithms on small sets of real data, it makes sense to move directly from conceptualization to realization. By having a real implementation we can test, experiment with, and improve on, we can speed up the process of converging on a final solution. Without a working implementation, we cannot infer anything about a given technique, or its applicability to our problem, rendering subsequent objectives useless.

Task I: Implementation

We will use the MATLAB programming language to prototype algorithms. MATLAB is a good choice for a variety of reasons: every team member has at least a basic familiarity with it from prior coursework. It is also distributed with a sizeable selection of machine learning and computer vision algorithms already implemented, drastically reducing the overhead required to get an implementation working. Finally, MATLAB scripts are relatively easy to modify and debug.

Measuring Completion

We anticipate that the pursuit of perfection will interfere with the completion of a working, if not perfect implementation. It will be essential that each implementation is completed, even if it becomes apparent that it is not working effectively. The main reason for this is to ensure that we can compare each technique with the others on a level field. But an even more important reason is that an algorithm which appears to be the loser may actually outperform another technique with slight modifications. An implementation is complete when it successfully executes and produces a reasonable result, even if that result is not perfect.

Objective III

Verify efficacy of chosen algorithm(s)

Rationale

Without accurate and repeatable results, any form of diagnosis is nearly useless. Therefore, it is essential that our final solution meets our expectations for sensitivity, specificity and error in parasite density, as well as computational time on low power hardware.

Task I: Collect Algorithm Results

The Malaria Research and Reference Reagent Resource Center (MR4) provides sample slides of stained blood smears along with information about the quantity and identification of the malarial parasites present. A microscope and camera will be used to capture digital images of the sample slides provided by MR4. The chosen algorithms will then be evaluated using the digital images. The results of the analysis of the algorithms on the MR4 data set will then be analyzed.

Measuring Completion

Metrics of performance include sensitivity and specificity. The algorithm must be sensitive enough to determine whether a sample has malaria present or not in at least 90% of the cases studied. The algorithm must also yield specific results such that at least 90% of the samples where malaria is detected actually have malaria. Therefore, our algorithm must have at least 90% specificity and 90% sensitivity.

The parasite density results obtained by our algorithm can then be plotted against the data given by the MR4 samples. The R^2 value of the regression fit will enable us to determine how our measured results compare to the accepted parasite density values given by MR4. The R^2 of this correlation must be greater than 0.9.

Objective IV

Test final algorithm on a real data set

Rationale

Though we can test our algorithms extensively in controlled lab environments, our end product will be not be effective if it does perform well in clinical trials. The obvious final step then is to deploy our proposed device in clinics and gather the results.

Task I: Develop Device

Develop a device to work with the algorithms we previously developed. This will largely involve programming microcontrollers and/or small computers to implement our algorithm using images taken from an integrated microscope. This hardware of this device will be simple compared to the software contained within it. We plan on prototyping and designing the device using tools available at Georgia Tech.

Task II: Test Device

Work with the CDC to test the device on real data sets. The CDC has access to both malarial blood and the resources required to do clinical trials. We plan to stay in contact with the malaria doctors at the CDC who will likely have the connections to help us accomplish this task.

Task III: Analyze Device Results

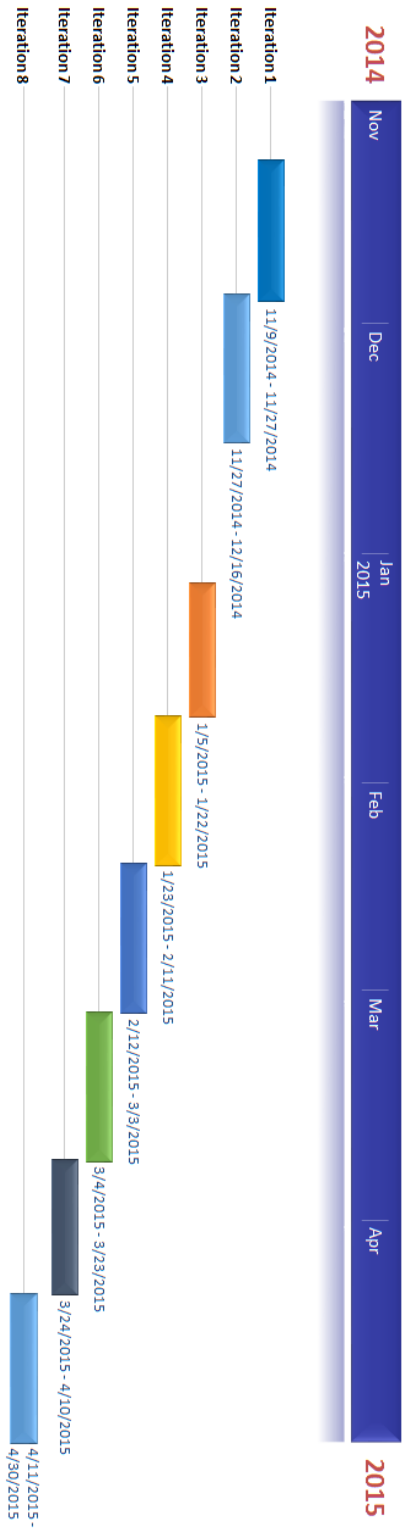
We anticipate using our device to analyze a patient's blood smear, while a simultaneous analysis is performed with at least two microscope technicians. The results from our device can then be compared to the standard human-performed analysis. This method of testing will be more expensive and resource intensive than comparing the device results to one human-performed test, but should help resolve any variability in human-performed analysis.

Measuring Completion

The metrics of performance will be similar to those described in Objective III. Again, the main metrics will be sensitivity and specificity. The device algorithm must have at least 90% specificity and 90% sensitivity. The algorithm results will be compared against the standard microscopy results.

The parasite density results obtained by our algorithm can then be plotted against the average density results given by the human technicians. The R^2 value of the regression fit will enable us to determine how our measured results compare to the accepted parasite density values given by human technicians. The R^2 of this correlation must be greater than 0.9.

Timeline



This timeline represents an Agile development methodology. Each iteration will last two weeks, and represent one cycle of the three main objectives enumerated above.

Budget

ITEM	SEMESTER	QTY	PRICE	TOTAL PRICE
4'X8' MDF Board	Spring 2015	1	$ 60.00	$ 60.00
AmScope M148C Compound Monocular Microscope,	Spring 2015	1	$ 76.83	$ 76.83
AmScope MD35A	Spring 2015	1	$ 39.98	$ 39.98
Raspberry Pi Model B+	Spring 2015	1	$ 38.09	$ 38.09
Travel to CDC	Spring 2015	3	$ 30.00	$ 90.00
Arduino	Spring 2015	1	$ 24.95	$ 24.95
Grand Total				**$329.85**

References

[1] Naval Medical Research Center. *Malaria Research.* Retrieved from http://www.med.navy.mil/sites/nmrc/Pages/id_m.htm

[2] World Health Organization. (Oct. 2011). *Eliminating Malaria: Learning From the Past, Looking Ahead.* Retrieved from: http://www.rbm.who.int/ProgressImpactSeries/docs/report9-en.pdf

[3] World Health Organization. (2014). *Malaria*. Retrieved from http://www.who.int/mediacentre/factsheets/fs094/en/

[4] Arguin, P. (2014, September 9). Personal interview.

[5] Msellem, M.I., et al. (2009). Influence of Rapid Malaria Diagnostic Tests on Treatment and Health Outcome in Fever Patients, Zanzibar—A Crossover Validation Study. *PLOS Medicine*. doi: 10.1371/journal.pmed.1000070

[6] F.B. Tek, A.G. Dempster, and I. Kale, "Computer Vision for Microscopy Diagnosis of Malaria," *Malaria Journal,* vol. 8, no. 153, July, 2009.

[7] Moody, A. (2002). Rapid Diagnostic Tests for Malaria Parasites. *American Society of Microbiology.* doi: 10.1128/CMR.15.1.66-78.2002

[8] C. Leighton and R. Foster, "Economic Impacts of Malaria in Kenya and Nigeria". Major Applied Research Paper 6, Health Financing and Sustainability (HFS) Project, Abt Associates Inc, 1993.