EDUCATIONAL GAME FOR MALARIA DIAGNOSIS

Project report submitted in complete fulfillment of the requirement
for the Degree of Bachelor of Technology

in

Computer Science and Engineering

By

Puneet Tyagi (121320)

Under the supervision of

Mr. Arvind Kumar

to

Department of Computer Science & Engineering and Information
Technology

**Jaypee University of Information Technology Waknaghat, Solan -
173234, Himachal Pradesh**

Candidate's Declaration

I hereby declare that the work presented in this report entitled **"Educational Game For Malaria Diagnosis"** in complete fulfillment of the requirements for the award of the degree of **Bachelor of Technology** in **Computer Science and Engineering** submitted in the department of Computer Science & Engineering and Information Technology**,** Jaypee University of Information Technology, Waknaghat is an authentic record of my own work carried out over a period from August 2015 to May 2016 under the supervision of Mr. Arvind Kumar (Assistant Professor, CSE and IT Department).

The matter embodied in the report has not been submitted for the award of any other degree or diploma.

Puneet Tyagi (121320)

This is to certify that the above statement made by the candidate is true to the best of my knowledge.

Mr. Arvind Kumar

Assistant Professor

CSE Department

Dated:

Acknowledgement

I have made a lot of efforts in completion of this project. However, it would not have been possible without the kind support and help of many individuals. I would like to extend my sincere thanks to all of them.

I am highly indebted to Mr. Arvind Kumar for his guidance and constant supervision as well as providing me with all the necessary information regarding the project and also for his support in completing this project.

I would also like to express my gratitude towards the faculty of Jaypee University of Information Technology, Waknaghat for their co-operation and encouragement which helped us in completion of this project.

My thanks and appreciations also goes to every individual who have willingly helped us out in one way or the other.

Puneet Tyagi (121320)

Contents

S. No. **Topic** **Page No.**

1. Introduction

2. Literature Survey

3. System Development

4. Performance Analysis

5. Conclusion

List of Abbreviations

1. CAPTCHA - Completely Automated Public Turing test to tell Computers and Humans Apart
2. CS - Crowd sourcing
3. GWAP - Game With A Purpose
4. HBC - Human Based Computation
5. HTML - HyperText Markup Language
6. MAP - Maximum A Posteriori
7. PC - Personal Computer
8. OT - Optimization Techniques
9. RBC - Red Blood Cells
10. WBC - White Blood Cells
11. WHO - World Health Organization

List of Figures

Abstract

According to the latest WHO estimates, released in September 2015, there were about 214 million cases of malaria in the year 2015 and about 438,000 people were killed. The report also states that approximately 3.2 billion people – about half of the world's population – are at risk of malaria. Most malaria cases and deaths occur in sub-Saharan Africa. However, Asia, Latin America, and, to a lesser extent the Middle East and parts of Europe, are also at risk. Early diagnosis and treatment of malaria reduces disease and prevents deaths. It also contributes to reducing malaria transmission.

The standard practice for diagnosing malaria consists of detecting the parasites, then counting the number of parasites in a blood smear sample. A pathologist must typically check on the order of 1,000 individual red blood cells under a high-magnification light microscope before being able to reliably call a sample healthy or negative for malaria. This unfortunately, is a time-consuming and challenging task given the large number of cases [3].

We aim to develop a self-learning distributed image analysis and diagnostics framework that is based on crowd-sourced games.

Human Computation methods such as crowdsourcing and games with a purpose (GWAP) have each recently drawn considerable attention for their ability to synergize the strengths of people and technology to accomplish tasks that are challenging for either to do well alone [7].

Our goal is to use knowledge that comes from the analysis of non-diagnostic experts to help the diagnostic experts improve the way that they do their malaria analysis.

We will report the results and the analysis of this project towards diagnosis of malaria infected human red blood cells (RBCs), where binary responses from untrained individuals are combined together, resulting in an accuracy level that is comparable to those of expert medical professionals.

Chapter-1 INTRODUCTION

1.1 Human Computation

Human-based computation (HBC), is basically a computer science technique in which a machine performs its function by outsourcing certain steps to human beings. This approach uses differences in abilities and alternative costs between humans and computer agents to achieve human-computer interaction.

In traditional computation, a human gives a problem to computer to solve; a human provides an overall problem description and an algorithm to the computer, and receives a final solution which needs to be interpreted. Human-based computation frequently reverses the roles; the computer asks a person or a large group of people to solve a problem, then collects, interprets, and combines their solutions.

Although computers have advanced significantly in many respects over the last 50 years, they still cannot compete with the basic conceptual intelligence or perceptual capabilities that most humans take for granted. By leveraging human skills and abilities in a new way, we hope to solve large-scale computational problems that computers cannot yet solve and begin to teach computers many of these human talents.

Online games constitute a general mechanism for using brain power to solve open computational problems. Each problem requires the careful design of a game developed to be enjoyable and, at the same time, guarantee that game-play correctly solves instances of the problem [3].

Instead of using a silicon processor, these processes run on a processor consisting of ordinary humans interacting with computers over the Internet. These games are referred as “human algorithm games [3]”.

Another example of the general approach for utilizing human processing power is CAPTCHA, an automated test that humans can pass but that current computer programs cannot. CAPTCHAs take advantage of the power of human cycles in order to differentiate people from computer. The ability to do this has important applications in practical security.

Human Computation organizes workers through a task market with APIs, and software-as-a-service protocols that allow employers / requesters to receive data generated by workers directly in to IT systems. As a result, many employers attempt to manage workers automatically through algorithms rather than responding to workers on a case-by-case basis or addressing their concerns. Responding to workers is difficult to scale to the employment levels enabled by human computation microwork platforms.

1.2 Crowd Sourcing

Crowdsourcing has been defined as the act of taking a job traditionally performed by a designated agent (such as an employee or a student) and making it available to an undefined, generally large group of persons in the form of an open problem.

Although crowdsourcing is not a new concept, the recent rise in attention placed on crowdsourcing is due to several factors, including the ubiquity of the Internet, the need to increase performance on tasks that computers cannot do well (such as relevance judgments, geo-tagging, and image annotations).

Online games offer a well-established form of entertainment; demographic studies show that the number of hours that people participate in games per week is increasing. Game play is no longer limited to a specific player demographic, and games have been gaining widespread acceptance as a tool to accomplish a specific goal, such as a supplement to learning. It is becoming easier to cajole players to accomplish a task while being entertained, which has been behind much of the recent development and use of games [6].

GWAP also have the ability to make mundane tasks engaging, provide a lower cost-per-task than even crowdsourcing can claim, and, which compared with crowdsourcing, has shown the potential of reducing noisy inputs [6].

We describe a crowd-sourcing based solution for handling large quantities of data that are created using next generation lab-on-a-chip platforms. We show that in cases where the diagnosis is a binary decision (e.g., positive vs. negative, or infected vs. uninfected), it is possible to make correct diagnosis by crowd-sourcing the basic given data (e.g., microscopic images of specimens/cells) using entertaining digital games that are played on PCs or tablets.

1.3 Problem Statement

In medical science, when the diagnosis is a binary decision (e.g., positive versus negative, or infected vs. uninfected), it is possible to make accurate diagnosis by crowd-sourcing the raw data. The general idea behind this project is the crowd sourcing bio medical image analysis and related diagnoses. This crowd sourced analysis can be especially powerful if we have the history of each expert's error rates and large number of people are involved. In this game, we have gone for malaria because it is a disease that still afflicts a large number of people around the globe, and is most prevalent in impoverished locations of the world.

1.4 Objectives

To design a malaria-diagnosis gaming platform where the non-expert gamers can collectively come close to improve the accuracy of an expert diagnostician. The game can also be used for training individuals and giving them information about malaria and how it can be detected and cured, especially in developing countries where medical facilities are not of good standards.

1.5 Methodology

In this highly educational game, all the images of blood cells are acquired, then these images are pre-processed. Upon preprocessing, these images are stored in database. When the game begins, the player records the response for each image, these responses are then stored and the final result or score is generated. The responses generated can be combined together to get results on some cases which are difficult to diagnose.

1.6 Organization

In Chapter 1, we have discussed about human computation and crowd sourcing. These are the concepts that guides the entire project and are used quite often to explain certain things.

In Chapter 2, we have discussed about various research papers and journals that we have gone through in preparation of this project.

In Chapter 3, we have discussed about how we have developed the project, its basic design, algorithm used and various figures depicting the working of project.

In Chapter 4, we have discussed about how we have measured the performance of the project, the different types of testing used and how we have obtained results and their significance.

In Chapter 5, we have concluded the project, its future scope and various applications.

Chapter -2 LITERATURE SURVEY

2.1 Overview of Research Papers/Journals

Title -	Optimization Techniques for Human Computation-enabled Data Processing Systems [9]
Author-	Adam Marcus
Date of Conference -	September 2012
Published In -	International Journal of Enterprise Computing Systems
Location -	Massachusetts Institute of Technology

Crowdsourced labor markets make it possible to recruit large numbers of people to complete small tasks that are difficult to automate on computers. These marketplaces are increasingly widely used, with projections of over $1 billion being transferred between crowd employers and crowd workers by the end of 2012. While crowdsourcing enables forms of computation that artificial intelligence has not yet achieved, it also presents crowd work designers with a series of challenges including describing tasks, pricing tasks, identifying and rewarding worker quality, dealing with incorrect responses, and integrating human computation into traditional programming frameworks.

In this dissertation, we explore the systems-building, operator design, and optimization challenges involved in building a crowd-powered work management system. We describe a system called Qurk that utilizes techniques from databases such as declarative work definition, high-latency work execution, and query optimization to aid crowd-powered work developers. We study how crowdsourcing can enhance the capabilities of traditional databases by evaluating how to implement basic database operators such as sorts and joins on datasets that could not have been processed using traditional computation frameworks. Finally, we explore the symbiotic relationship between the crowd and query optimization, enlisting crowd workers to perform selectivity estimation, a key component in optimizing complex crowd-powered work.

While there is great promise in the ways in which we can use crowds to extend the capabilities of computation, working with crowd workers has its own set of challenges. The biggest challenge is the Code reuse. Programmers encounter unnecessarily repeated code in two ways as they implement crowd-powered workflows.

The first is in implementing the logic behind common functionality, such as filtering or labeling a collection of items in a dataset, or performing entity resolution between two datasets.

The second is in the basic plumbing of crowd work: generating HTML tasks, assignment of workers to tasks, asynchronously collecting worker responses, verifying result quality across workers, and compensating the workers with good result quality.

To keep costs low, one can simply pay workers less to perform each task. This has the effect of increasing the time to complete a task, as workers prefer other tasks in the market. It can also reduce quality as workers increase the rate at which they perform the task, which may introduce new errors. Paying workers less also has ethical implications that I address at the end of this dissertation. Another approach to driving down cost, and sometimes even time, is to batch tasks together. Rather than labeling a single image as inappropriate, a worker might label two at a time, reducing the overhead of waiting for their work to be submitted or for the next task to load.

Workers can make mistakes or try to perform less accurate work to finish tasks faster. (Workers with consistently low-quality responses are sometimes called spammers.") Because of the uncertainty in worker responses, it is impossible to guarantee 100% data accuracy.

Title – Human Computation [2]

Author- Luis von Ahn

Date of Conference – December 7, 2005

Published In - National Science Foundation (NSF) grants

Location - Carnegie Mellon University, Pittsburgh

Tasks like image recognition are trivial for humans, but continue to challenge even the most sophisticated computer programs. This thesis introduces a paradigm for utilizing human processing power to solve problems that computers cannot yet solve. Traditional approaches to solving such problems focus on improving software. I advocate a novel approach: constructively channel human brainpower using computer games. For example, the ESP Game, introduced in this thesis, is an enjoyable online game — many people play over 40 hours a week — and when people play, they help label images on the Web with descriptive keywords. These keywords can be used to significantly improve the accuracy of image search. People play the game not because they want to help, but because they enjoy it.

There are three other examples of “games with a purpose”:
Peekaboom, which helps determine the location of objects in images, Phetch, which collects paragraph descriptions of arbitrary images to help accessibility of the Web, and Verbosity, which collects “common-sense” knowledge.

Then, there are CAPTCHAs, automated tests that humans can pass but computer programs cannot. CAPTCHAs take advantage of human processing power in order to differentiate humans from computers, an ability that has important applications in practice.

The results of this thesis are currently in use by hundreds of Web sites and companies around the world, and some of the games presented here have been played by over 100,000 people. Practical applications of this work include improvements in problems such as: image search, adult-content filtering, spam, commonsense reasoning, computer vision, accessibility, and security in general.

In this paradigm, we treat human brains as processors in a distributed system, each performing a small part of a massive computation. Unlike computer processors, however, humans require an incentive in order to become part of a collective computation. We propose online games as a means to encourage participation in the process.

We argue that games constitute a general mechanism for using brain power to solve open computational problems. Each problem requires the careful design of a game developed to be enjoyable and, at the same time, guarantee that game-play correctly solves instances of the problem. We argue that designing such games is much like designing computer algorithms: the game needs to be proven correct and enjoyable; its efficiency can be analyzed; more efficient games can supersede less efficient ones, etc. Instead of using a silicon processor, these “algorithms” run on a processor consisting of ordinary humans interacting with computers over the Internet.
We refer to these games as “human algorithm games.”

Title – Applying Human Computation methods to Information Science [4]

Author- Christopher Glenn Harris

Date of Conference – December 2013

Published In - Iowa Research Online

Location - University of Iowa

Human Computation methods such as crowdsourcing and games with a purpose (GWAP) have each recently drawn considerable attention for their ability to synergize the strengths of people and technology to accomplish tasks that are challenging for either to do well alone. Despite this increased attention, much of this transformation has been focused on a few selected areas of information science.

This thesis contributes to the field of human computation as it applies to areas of information science, particularly information retrieval (IR). We begin by discussing the merits and limitations of applying crowdsourcing and game-based approaches to information science. We then develop a framework that examines the value of using crowdsourcing and game mechanisms to each step of an IR model.

We identify three areas of the IR model that our framework indicates are likely to benefit from the application of human computation methods: acronym identification and resolution, relevance assessment, and query formulation. We conduct experiments that employ human computation methods, evaluate the benefits of these methods and report our findings. We conclude that employing human computation methods such as crowdsourcing and games, can improve the accuracy of many tasks currently being done by machine methods alone. We demonstrate that the best results can be achieved when human computation methods augment computer-based IR processes, providing an extra level of skills, abilities, and knowledge that computers cannot easily replicate.

Crowdsourcing, as a human computation approach, has been defined as the act of taking a job traditionally performed by a designated agent (such as an employee or a student) and making it available to an undefined, generally large group of people in the form of an open call (Howe, 2006). Although crowdsourcing is not a new concept, the recent rise in attention placed on crowdsourcing is due to several factors, including the ubiquity of the Internet, the need to increase performance on tasks that computers cannot do well (such as relevance judgments, geo-tagging, and image annotations), the improved worldwide reach of micropayment methods, the cost of hiring experts, as well as the disparity of global economic labor demand and tight local labor restrictions. In fact, the large worker supply, little regulation, and low labor costs provide crowdsourcing’s strongest advantages.

As with many new technologies, the early days of crowdsourcing have primarily focused on the areas with the greatest need: repetitive, single-purpose tasks designed around a single objective, such as image classification, video annotation, form-based data entry, optical character recognition, translation, and document proofreading. However, as crowdsourcing begins to mature, it has begun a transformation, creating fascinating new opportunities for leveraging real-time human computation for a range of diverse and complex tasks, such as providing quality assurance in the peer review process in biological research, providing a detailed check of submitted expense receipts by the U.K.’s Members of Parliament, or analyzing spectrograms and sounds of whales in an attempt to decipher them.

Title – BioGames: A Platform for Crowd-Sourced Biomedical Image Analysis and Telediagnosis [3]

Authors- Sam Mavandadi, Steve Feng, Frank Yu, Stoyan Dimitrov, Richard Yu, and Aydogan Ozcan

Date of Conference – September 2012

Published In - GAMES FOR HEALTH JOURNAL: Research, Development, and Clinical Applications

Location - University of California, Los Angeles

Medical imaging has gone through a co-evolution along with the computer industry over the past three decades, with each medical imaging modality benefiting in major ways from the ever-increasing abilities of modern computers. The possibility to capture, store, and manipulate images digitally has brought upon a new age of medical imaging with a significant shift in focus toward more complex analysis software. Through shear computation and clever mathematical algorithms, modern medical imaging devices are capable of producing higher-quality images faster while exposing patients to much less harmful radiation. As part of this trend, an emerging field where we have focused most of our own efforts over the past few years is that of computational microscopy.

Another dimension of medical imaging's evolution has been a consequence of rapid advances in telecommunications and the coming of age of the Internet. These days an X-ray or a microscope slide image can be viewed almost instantaneously thousands of miles away from the point of capture by an expert who had no involvement in the imaging procedure. This unprecedented level of access to medical images and data is now opening up new approaches to medical diagnosis, heralding the age of telemedicine, where one can outsource medical diagnosis to doctors in faraway locations, while making it significantly easier to get a second opinion on a particular diagnosis.

For diagnosis of malaria, conventional light microscopy remains as one of the gold standard methods, with 165 million cases having been diagnosed through this method in 2010. A pathologist must typically check on the order of 1,000 individual red blood cells under a high-magnification light microscope before being able to reliably call a sample healthy or negative. This, unfortunately, is a time-consuming and challenging task given the large number of cases observed, resulting in a false-positive rate of, for example, approximately 60 percent in some developing countries. Such a high false-positive rate can lead to unnecessary treatments and hospitalizations.

To test our idea, we started by creating entertaining digital games (termed ‘‘BioGames’’) where the players were presented with a set of red blood cell images taken from potentially infected samples. They were allowed to choose to digitally ‘‘kill’’ or ‘‘bank’’ the infected and healthy cells, respectively. To be able to later combine the information generated by multiple gamers, we had to know how they were doing in terms of diagnosis accuracy as they went along playing the game. Toward this end, by carefully embedding some known images (i.e., control images) in our games, we were able both to assign scores to gamer performances and to quantify how well they could diagnose individual cells. Once the responses from all the gamers were collected, we could then combine them using techniques borrowed from telecommunications theory (by mathematically treating each gamer as a noisy telecommunication channel) to yield much more reliable diagnoses for the unknown blood cell images.

To quantify the performance of our gamers and our fusion algorithms, we also asked trained medical experts to individually check and label all of the images in our database, creating a gold standard label for each cell.

Looking forward, we believe that the ‘‘BioGames’’ platform can be extended to other biomedical image analysis and diagnosis problems, such as microscopic analysis of Papsmears for diagnosis of, for example, cervical cancer. Furthermore, in addition to binary diagnostic decisions, it can also accommodate a wider range of non-binary diagnostic possibilities. A component of this platform that requires further work, and is crucial for its success and wide-scale deployment, is the game itself. In order to attract and keep the interest of gamers, highly entertaining games that can be played on a multitude of platforms such as PCs, mobile phones, tablets, and other gaming devices need to be designed/created. As such, a possible avenue that is worth exploring toward continuous development of entertaining games is to open up the ‘‘BioGames’’ platform to developers around the world and allow for a game-developer community to form.

In summary, we believe that such innovative uses of digital gaming technologies for telemedicine applications could in the very near future open up new avenues for delivering faster, more accurate, and cost-effective diagnosis to the masses globally, as well as for wide-scale and efficient training of medical professionals and diagnosticians.

Title -	Crowd-sourced BioGames: managing the big data problem for next generation lab-on-a-chip platforms [1]
Authors-	Sam Mavandadi
Date of Conference –	August 2012
Published In -	The Royal Society of Chemistry
Location -	University of California, Los Angeles

We describe a crowd-sourcing based solution for handling large quantities of data that are created by e.g., emerging digital imaging and sensing devices, including next generation lab-on-a-chip platforms.

We show that in cases where the diagnosis is a binary decision (e.g., positive vs. negative, or infected vs. uninfected), it is possible to make accurate diagnosis by crowd-sourcing the raw data (e.g. microscopic images of specimens/cells) using entertaining digital games (i.e., BioGames) that are played on PCs, tablets or mobile phones.

We report the results and the analysis of a large-scale public BioGames experiment toward diagnosis of malaria infected human red blood cells (RBCs), where binary responses from approximately 1000 untrained individuals from more than 60 different countries are combined together (corresponding to more than 1 million cell diagnoses), resulting in an accuracy level that is comparable to those of expert medical professionals.

This BioGames platform holds promise toward cost-effective and accurate tele-pathology, improved training of medical personnel, and can also be used to manage the ‘‘Big Data’’ problem that is emerging through next generation digital lab-on-a-chip devices.

The internet revolution can be traced back to the invention of the transistor. It was the integration of transistors into electronic chips, and subsequently into computational systems that led to the development of computers and the internet.

For over three decades, the number of transistors in integrated circuits has been doubling approximately every 18 months, a trend that is dubbed as the Moore’s Law. Quite interestingly, a similar trend seems to exist for the number of pixels installed on mobile phone cameras. In other words the megapixel count of mobile phones has been following Moore’s Law since the wide-scale introduction of camera phones in 2002 with 0.3 megapixels, reaching 41 megapixels in 2012.

The massive volume of these consumer electronics components, reaching billions, has brought a phenomenal reduction in cost and unprecedented levels of access to such advanced digital devices despite their sophisticated hardware and software capabilities.

As a result of this revolution in digital electronics, we are now seeing a paradigm shift in imaging and sensing technologies with ultra-portable, cost-effective and high-throughput lab-on-a-chip platforms providing innovative solutions for e.g., point-of-care and telemedicine applications, among others.

At the centre of the BioGames platform lies crowd-sourcing, which has recently emerged as a powerful strategy for tackling computationally difficult scientific problems.

The basic idea behind a crowd-sourcing platform is to first break the task of interest into smaller pieces that can each be completed in a relatively short amount of time; and second to distribute these pieces of the problem to individual humans through a convenient medium, such as the internet; and finally to combine the collective responses of the individual participants to yield an optimal solution. A promising approach to distributing the data and finding human volunteers has been the use of entertaining computer games. In this approach the scientific or computational problem is embedded into a digital game and then distributed to gamers. The individual gamers may then cooperate, compete, or play independently to solve parts or the entire problem of interest.

Looking forward, the BioGames concept could help us better manage the Big Data problem emerging with the introduction of next generation imaging, sensing and lab-on-a-chip devices which all benefit from ubiquitous digital communications technologies (e.g., mobile phones, tablet PCs, etc) for wide-scale generation of massive amounts of biomedical and environmental data, even in resource poor settings and remote locations. Through intelligent crowd-sourcing and digital gaming strategies, we can potentially harness the power of human crowds and their innate visual pattern recognition and learning abilities to better sort, classify, diagnose, and manage this emerging Big Data.

In conclusion, here we described how the task of binary image-based diagnosis can be crowd-sourced to minimally trained individuals and yet yield accurate results. We reckon that under many circumstances this methodology is more practical than the use of automated computer algorithms for the same purpose since the human visual system offers a very low-cost and highly superior pattern recognition platform for such image understanding tasks.

Chapter-3 SYSTEM DEVELOPMENT

3.1 Introduction

In this project, we investigate whether the innate visual recognition and learning capabilities of untrained humans can be used in conducting reliable microscopic analysis of biomedical samples toward diagnosis. For this purpose, we designed entertaining digital games that are interfaced with artificial learning and processing back-ends to demonstrate that in the case of binary medical diagnostics decisions (e.g., infected vs. uninfected), with the use of crowd-sourced games it is possible to approach the accuracy of medical experts in making such diagnoses. Specifically, using non-expert gamers we report diagnosis of malaria infected red blood cells with an accuracy that is within 1.25% of the diagnostics decisions made by a trained medical professional.

Digital games have been used as effective means to engage an individual’s attention to computational tasks of interest. We take a similar strategy and demonstrate a platform to use digital gaming and machine learning to crowdsource the analysis of optical microscopy images of biomedical specimens through engaging the interest of human game players (i.e., gamers). The primary goal of this methodology is to accurately diagnose medical conditions, approaching the overall accuracy of medical experts, while only using non-expert gamers. The same method can also function as a telemedicine platform, where trained medical experts could be made part of our gamer crowd through various incentives.

Crowd-sourcing of microscopic analysis and related diagnosis through gaming is rather timely in several ways.

First, with rapid advances in mobile telecommunication and internet technologies such as mobile-phones, tablet PCs, etc., we have hundreds of millions of active users and potential gamers in the cloud that are all connected to a global network.

In addition to this massive crowd volume, over the last few years, there has been a significant effort to create cost-effective, compact and lightweight microscope designs such that even mobile-phones could be converted into microscopic analysis tools.

Fig 3.1.1 Biomedical data (e.g., images of thin blood smear samples) from individual light microscopes all around the world are transmitted to data centres where they are pre-processed and digitally distributed among gamers, which in turn diagnose and transmit their responses back. These individual results of the gamers are then fused toward a final diagnosis, the result of which is transmitted back to the point-of-care or the clinic/hospital. In the map above, orange-coloured regions show locations where risk of contraction of malaria still exists.

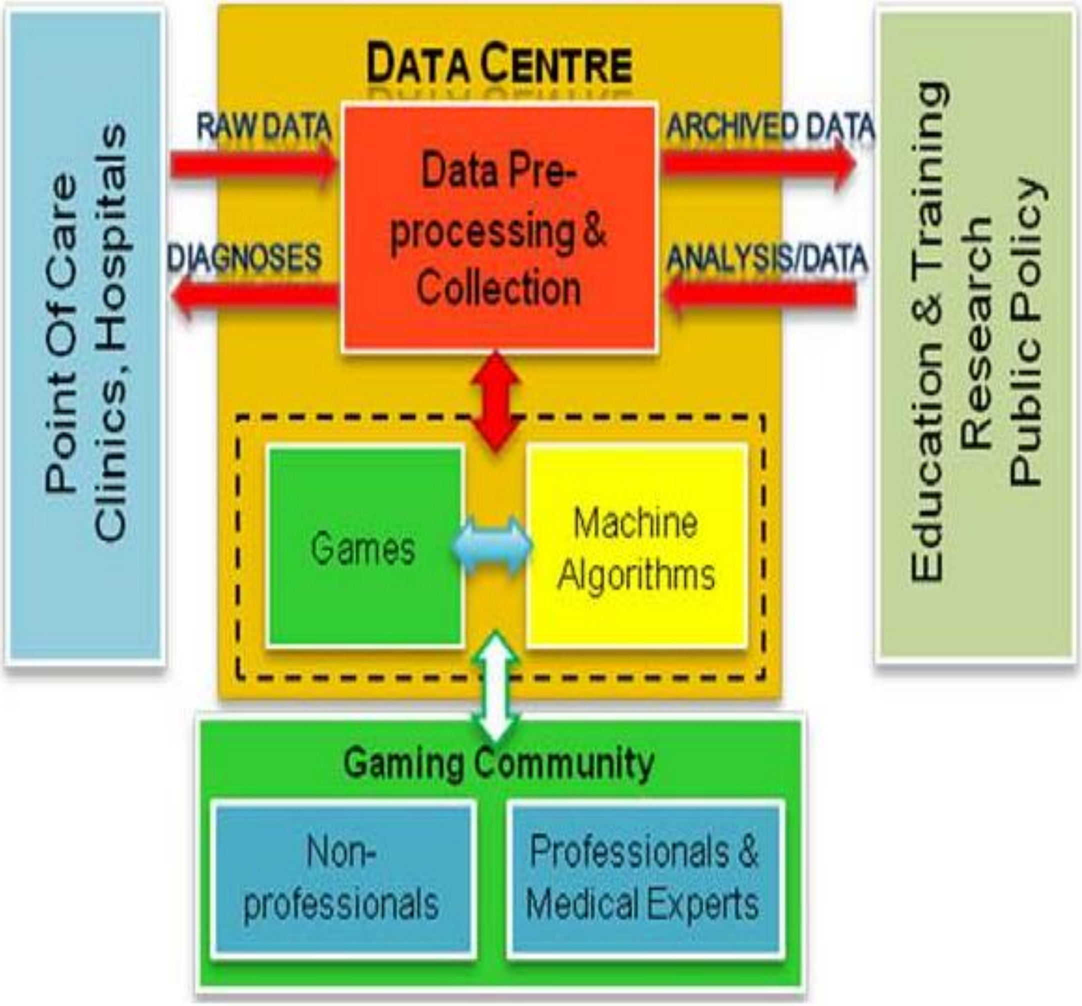


Fig 3.1.2 Block diagram of the presented platform.

To demonstrate the proof-of-concept of the above outlined framework, we chose malaria as the medical condition to be diagnosed, and developed a crowd-sourcing and distributed gaming platform that allows individuals from anywhere in the world to assist in identifying malaria infected red blood cells (RBCs) imaged under light microscopes.

For this initial demonstration, we chose malaria since it is still a major health problem in many tropical and sub-tropical climates, including much of sub-Saharan Africa (see Figure 3.1). It is the cause of, 20% of all childhood deaths in this region, and almost 40% of all hospitalizations in whole of Africa. For diagnosis of malaria, conventional light microscopy remains the gold standard method.

A pathologist must typically check between 100 and 300 different field-of-views (FOVs) of a thin blood smear (corresponding to inspection of at least 1,000 individual RBCs) using a light microscope with 100X objective lens before being able to reliably call a thin smear sample negative (i.e. not infected). This, however, is a very time-consuming task and a significant challenge given the large number of cases observed in these resource-poor settings. Furthermore, approximately 60% of the cases reported in sub-Saharan Africa are actually false-positives, leading to unnecessary treatments and hospitalizations.

3.2 Algorithm used

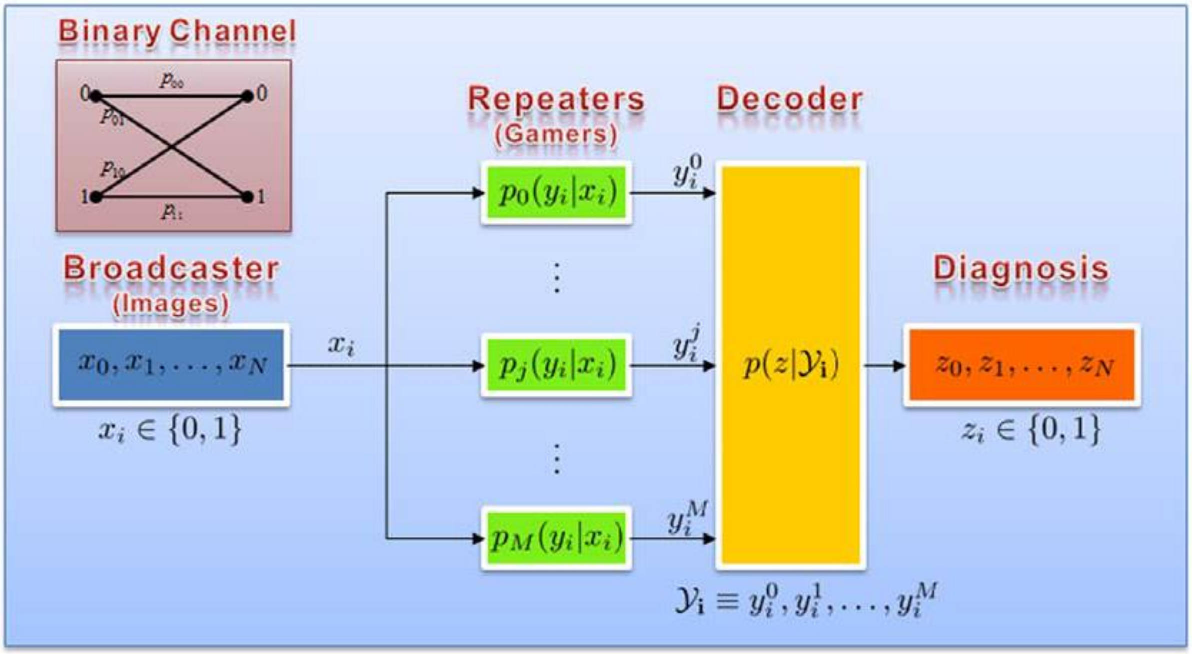


Fig 3.2.1 Overview of the gaming analysis framework.

The images are treated as a sequence of binary values that are broadcast by the server and are described by x_i having values 0 or 1. The sequence of symbols broadcasted to repeater units (i.e., the gamers within the crowd) that can be viewed as M parallel noisy channels. Our server acts as a broadcaster of a binary sequence and each gamer will act as a noisy Binary Channel, retransmitting the symbols back along with some errors In the most ideal case, the gamers will try to output the correct symbol for the inputs that they receive.

Using the control images, we compute the error probabilities associated with each gamer as p (yi |xi = l) where yi is the label that gamer k has provided for the ith cell image, xi is the true label of the image, and l is between 0 and 1. However, it is difficult to accurately estimate this asymmetric probability in our games due to the imbalance in the existence of positive and negative training data Each repeater transmits its own noisy version of the same input symbol to a decoder.

Similar to traditional communications systems, our broadcast unit can also include an encoder to increase the information redundancy prior to transmission to the repeaters/gamers. Given that the symbols being transmitted by the broadcast unit are not known a priori, the only appropriate coding scheme is the repetition code, where each symbol is repeated for an odd number of times prior to transmission. However, if the image is difficult to diagnose, then there is the chance that the gamer would not be consistent in making a decision

At the decoder, a majority vote is taken on the channel outputs. The decoder combines all the received repeater outputs and decodes a final output zi in the form of 0 and 1, which ideally will be the correct label or diagnosis for the input images. In designing the decoder for the gaming platform, we have taken a Maximum a Posteriori Probability (MAP) approach, which is used to obtain a point estimate of an unobserved quantity on the basis of empirical data.

3.3 Game Design

The game will be a program written in Java using the NetBeans IDE. We shall use the Java swings for containing our design elements.

Before starting to play the game, each gamer was given a brief explanation of the rules of the game and how malaria infected RBCs typically look with some example images.

To build a malaria infected RBC database, we used thin blood smear slides that contained mono-layers of cultured human RBCs which were infected by Plasmodium falciparum forms the source for our image dataset

After this, each gamer played a training game where she/he was required to successfully complete in order to continue playing the rest of the game. This test game consisted of various unique RBC images, where some of them were infected. The gamers were required to differentiate between the infected ones and the non-infected ones.

The gamers are then required to register their profile by entering all the required details. Upon registration, a unique user ID was assigned to each gamer and her/his individual diagnostics performance was tracked.

Furthermore, this game provided direct feedback to the players on their performance and their mistakes through a scoring mechanism. Since the labels (i.e., infected cell vs. healthy cell) of all the images were known a priori for the purposes of this training game, the player's score was updated throughout the game (i.e., positive score for correct diagnosis, and negative score for incorrect diagnosis).

It is important to note that there exists a large body of work on educational games. However, given that our focus was not to educate the players, and in fact it was to demonstrate the quality of diagnostic results that can be achieved through untrained (non-expert) individuals, this initial test/training game was designed in a simple repetitive fashion.

As the gamer goes through the game, he/she is presented with multiple frames of RBC images. The gamer has selects infected cells one by one till all the remaining cells in the current frame are designated. Within each frame, there are a certain number of cells whose labels (infected or healthy) are known to the game, but unknown to the gamers. These control cell images allow us to dynamically estimate the performance of the gamers (in terms of correct and incorrect diagnosis) as they go through each frame and also help us assign a score for every frame that they pass through. Once a frame is completed, a score is assigned based on the performance of the gamer only on the control images. The images and their order of appearance were identical among different gamers, thus allowing us to make a fair comparison among their relative performances.

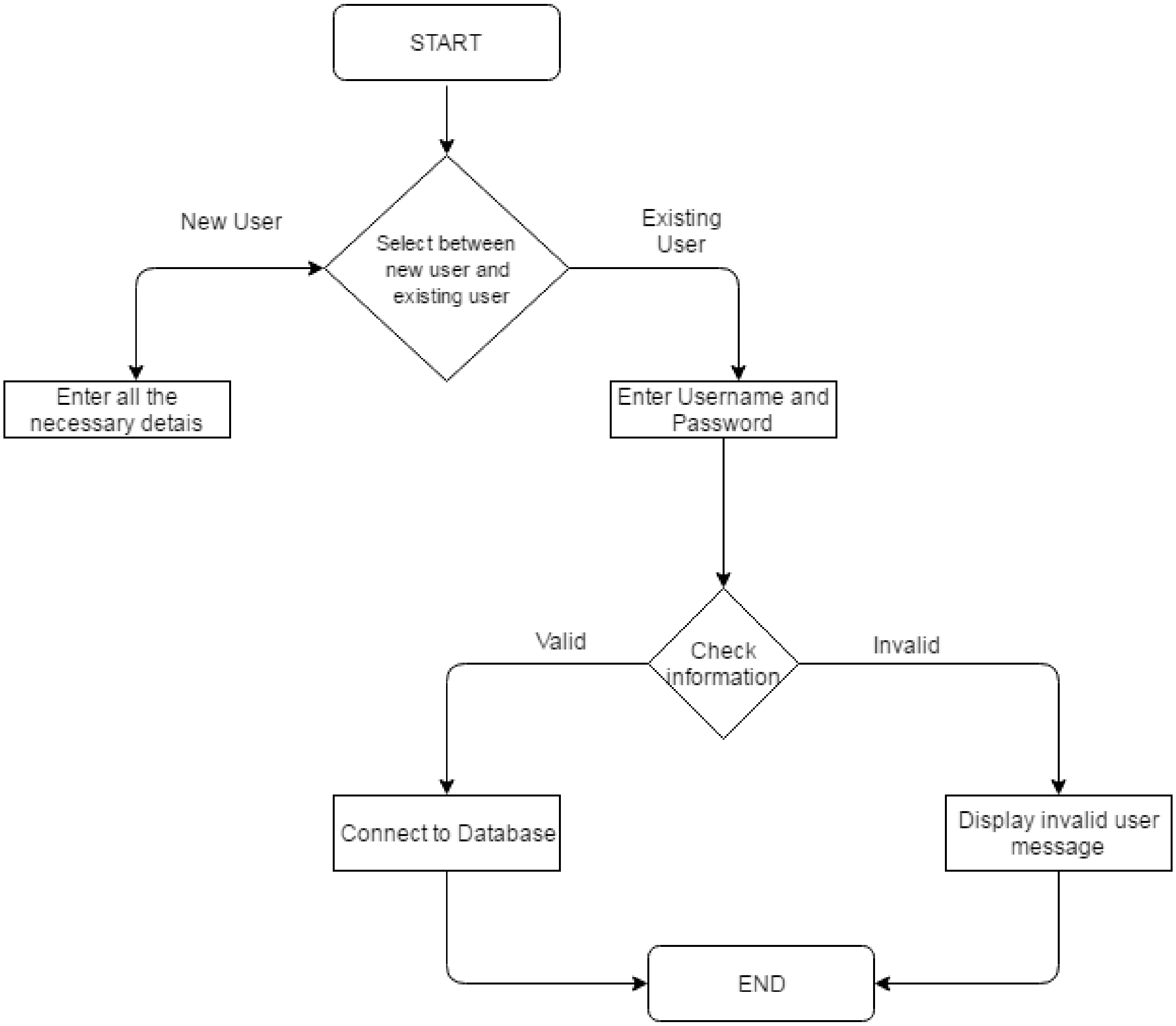


Fig 3.3.1 Login Page Flowchart

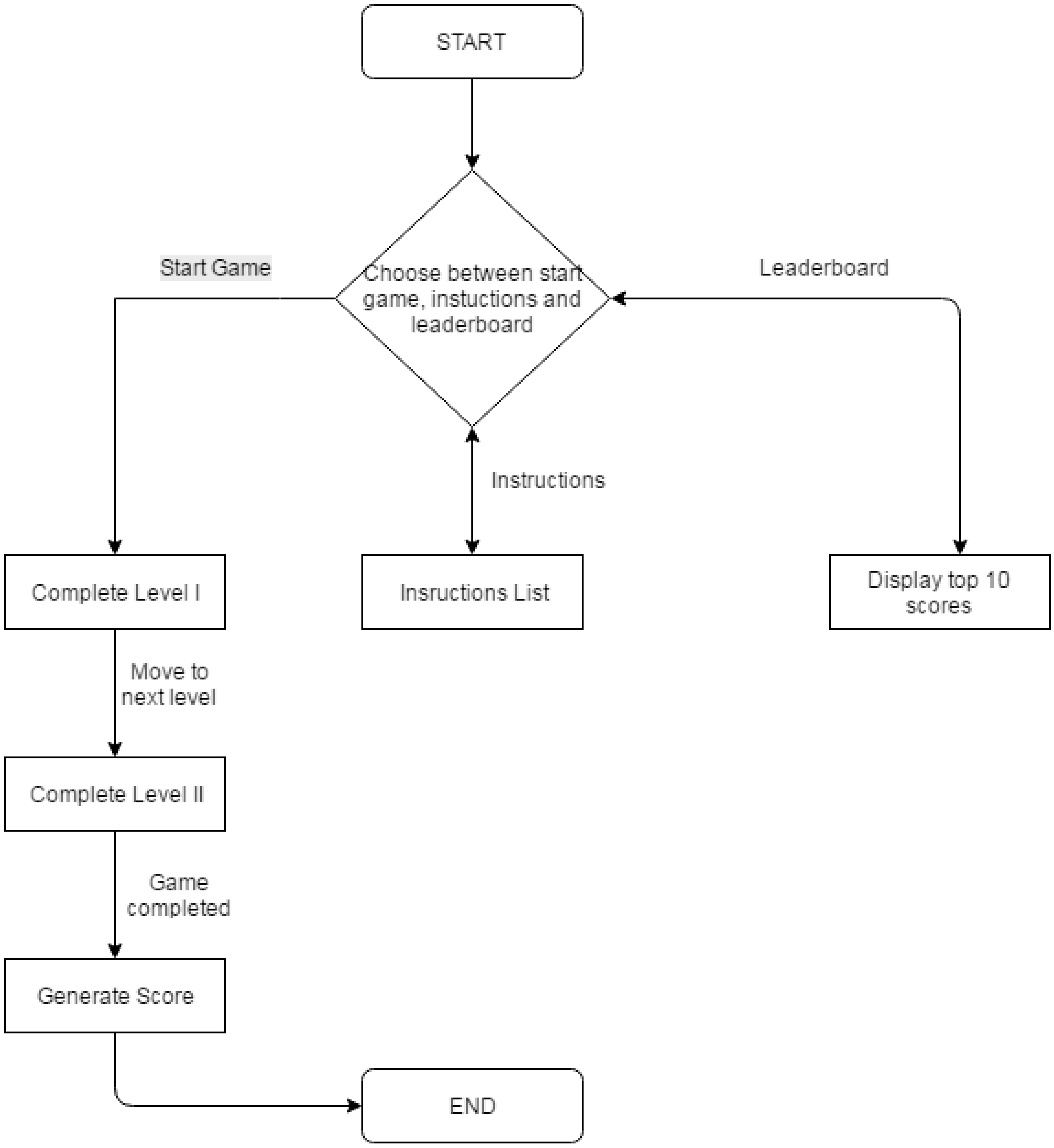


Fig 3.3.2 Gameplay Flowchart

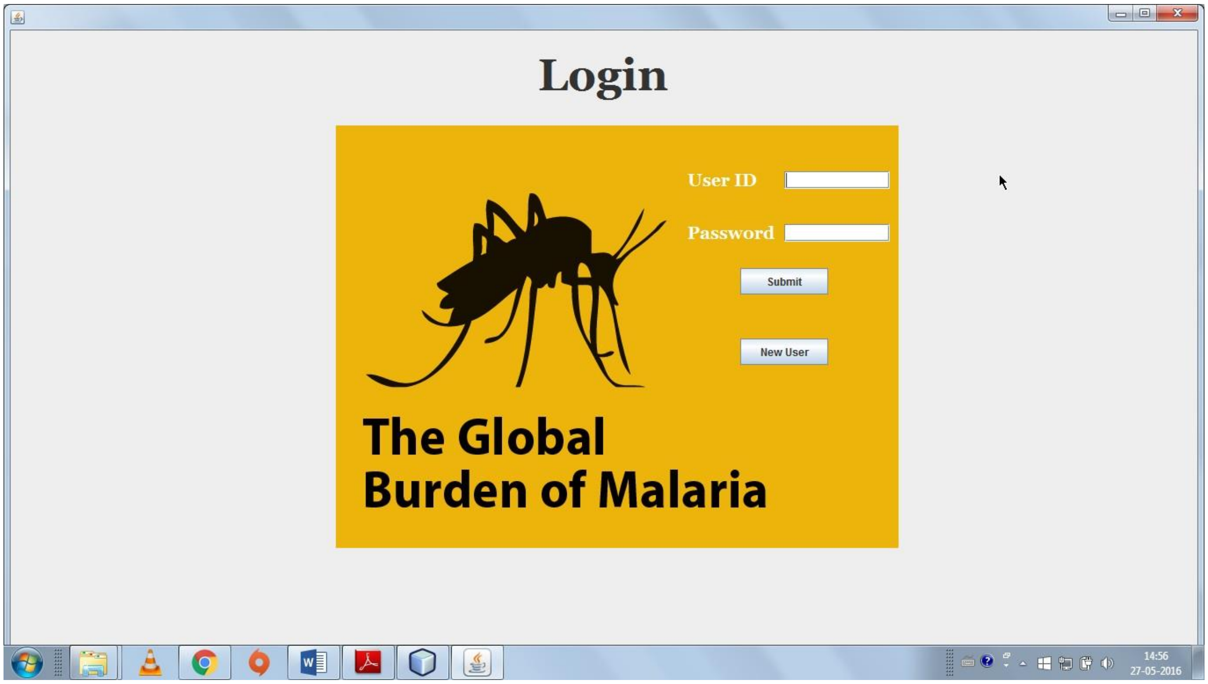
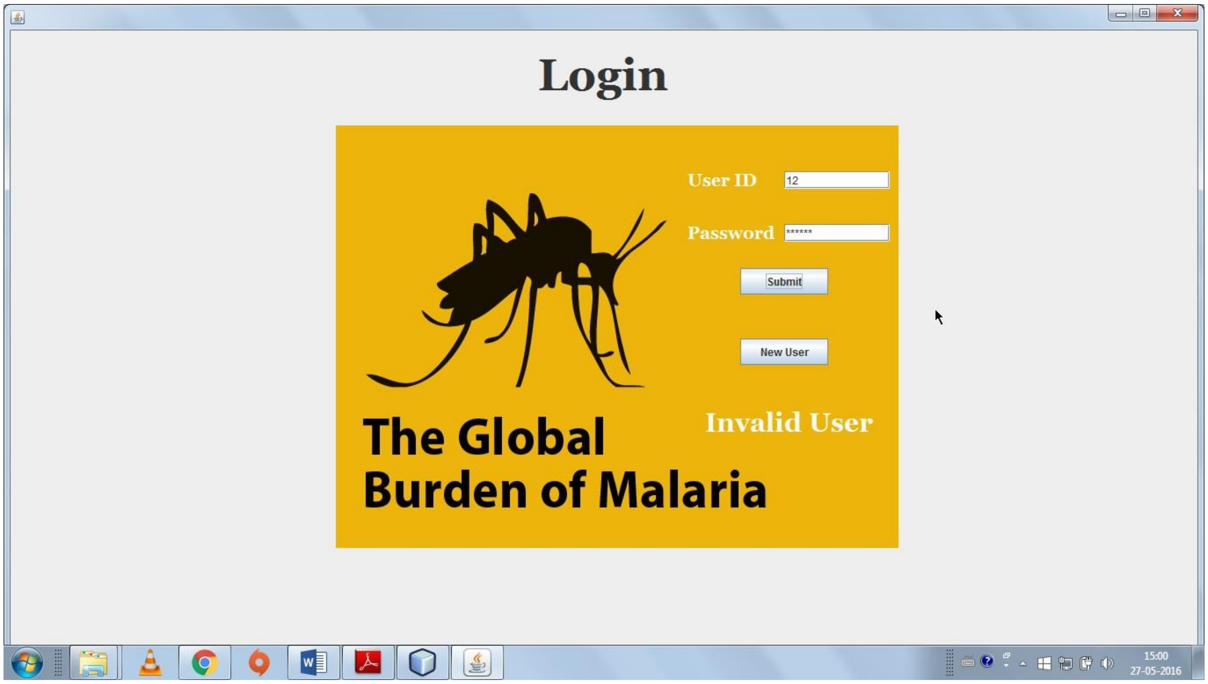


Fig 3.3.3 Login page

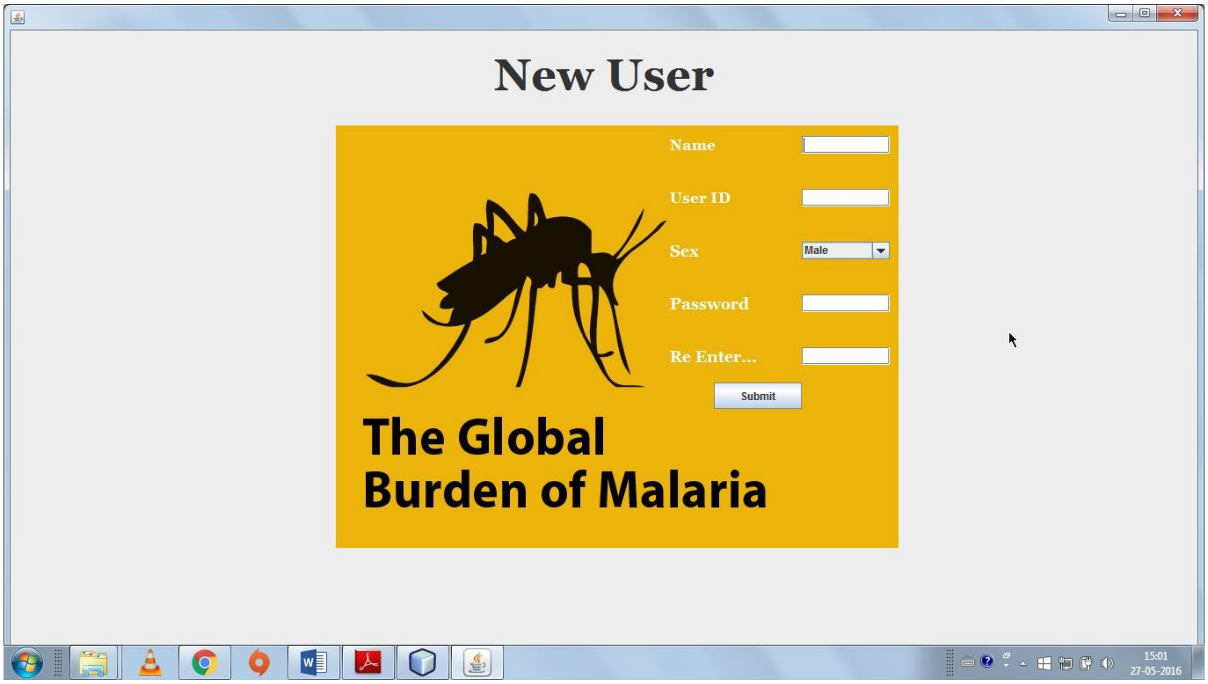


User ID 12

Password ******

Invalid User

Fig 3.3.4 If user enters invalid ID/ password



Name

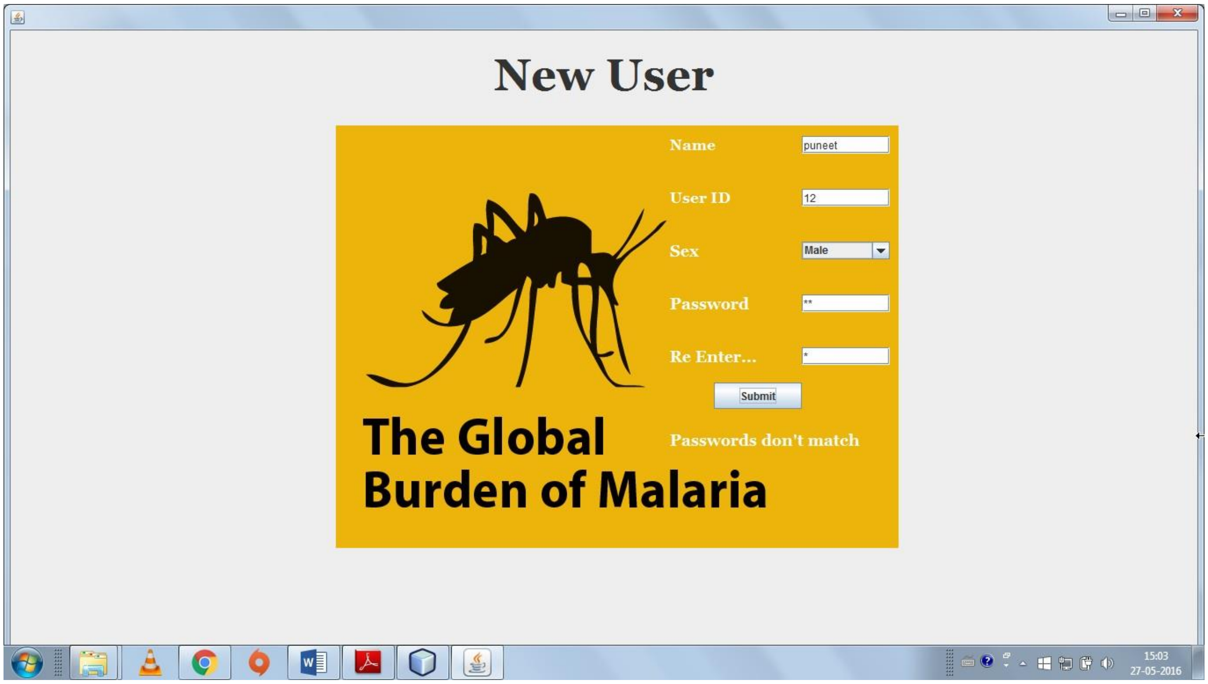
User ID

Sex: Male

Password

Re Enter...

Fig 3.3.5 For new user registration



Name: puneet

User ID: 12

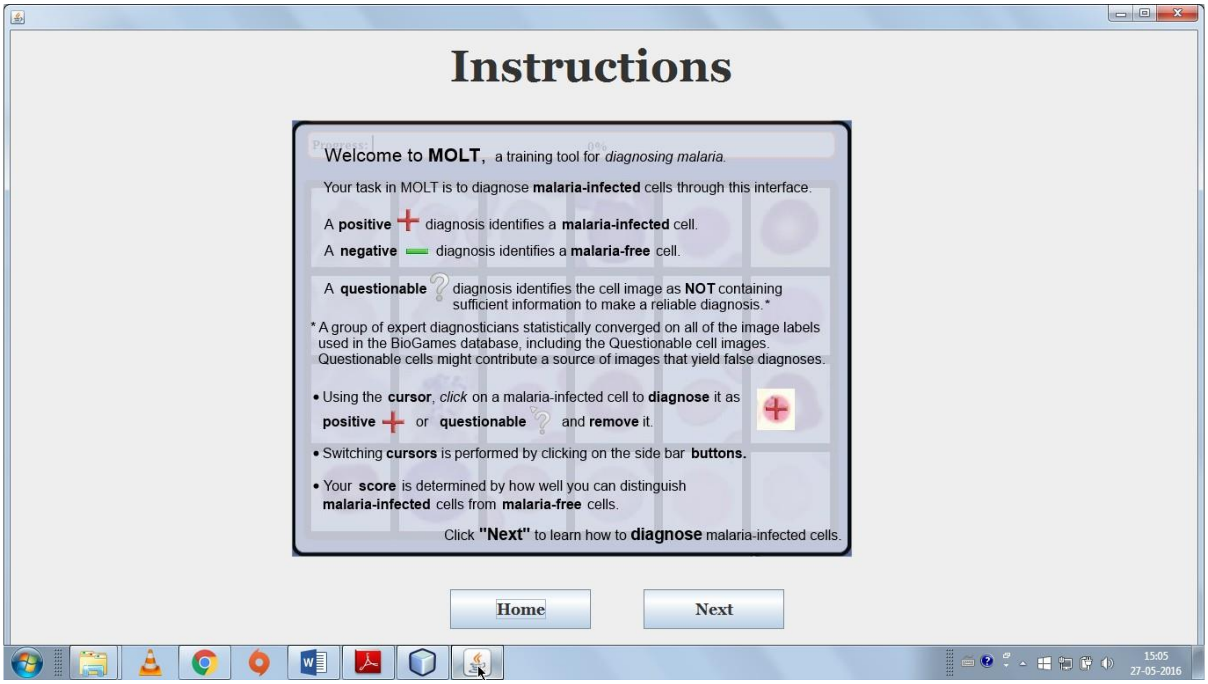
Sex: Male

Password: **

Re Enter...: *

Passwords don't match

Fig 3.3.6 If passwords entered doesn’t match



Welcome to **MOLT**, a training tool for *diagnosing malaria.*

Your task in MOLT is to diagnose **malaria-infected** cells through this interface.

A **positive** + diagnosis identifies a **malaria-infected** cell.

A **negative** — diagnosis identifies a **malaria-free** cell.

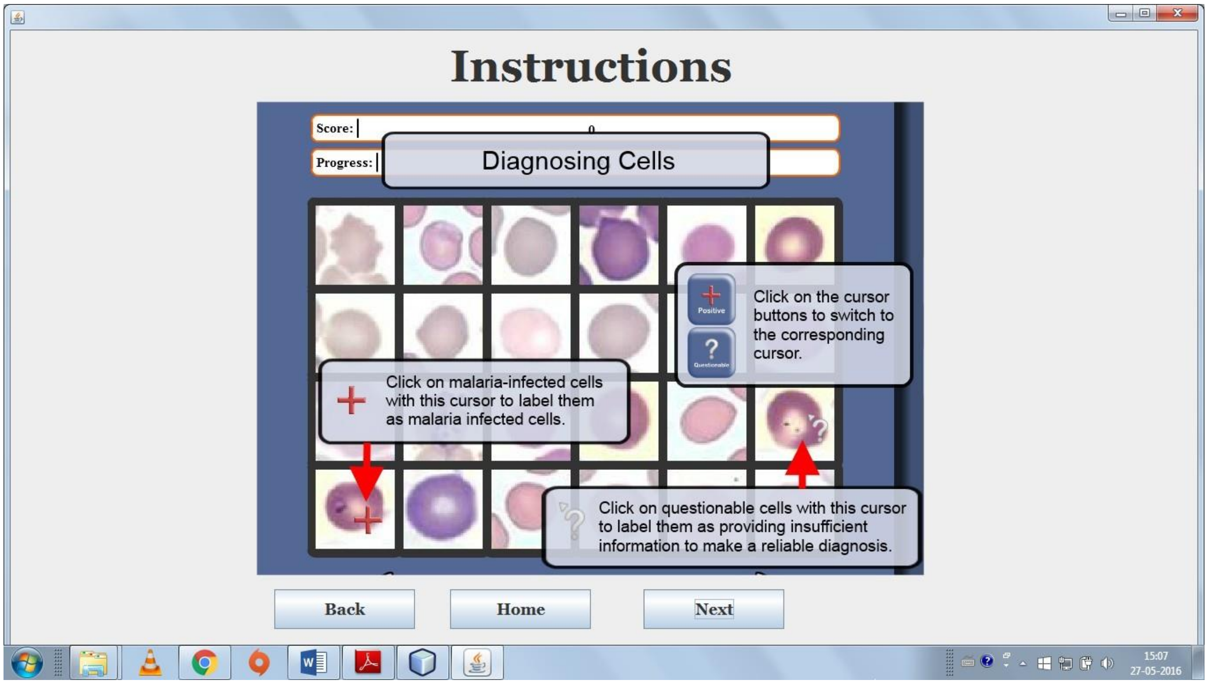
A **questionable** ? diagnosis identifies the cell image as **NOT** containing sufficient information to make a reliable diagnosis.*

* A group of expert diagnosticians statistically converged on all of the image labels used in the BioGames database, including the Questionable cell images. Questionable cells might contribute a source of images that yield false diagnoses.

- Using the **cursor**, *click* on a malaria-infected cell to **diagnose** it as **positive** + or **questionable** ? and **remove** it.
- Switching **cursors** is performed by clicking on the side bar **buttons.**
- Your **score** is determined by how well you can distinguish **malaria-infected** cells from **malaria-free** cells.

Click **"Next"** to learn how to **diagnose** malaria-infected cells.

Fig 3.3.7 Instructions set to play game

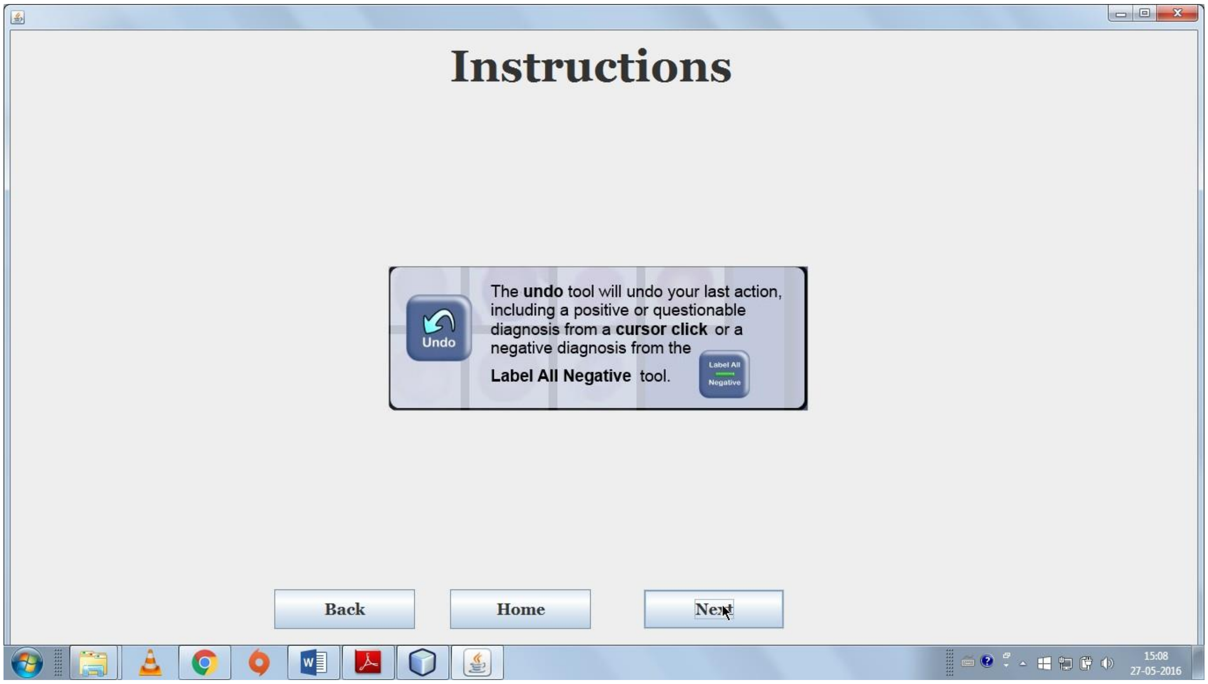


Click on the cursor buttons to switch to the corresponding cursor.

Click on malaria-infected cells with this cursor to label them as malaria infected cells.

Click on questionable cells with this cursor to label them as providing insufficient information to make a reliable diagnosis.

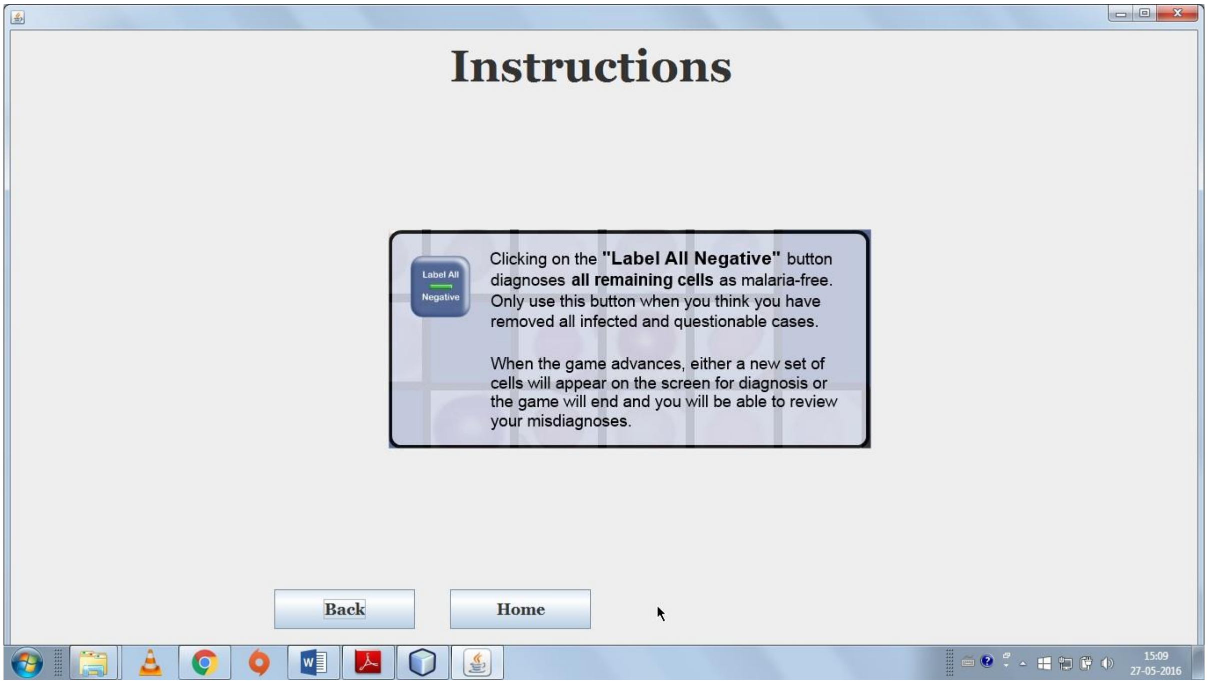
Fig 3.3.8 Instructions set to play game



Instructions

The **undo** tool will undo your last action, including a positive or questionable diagnosis from a **cursor click** or a negative diagnosis from the **Label All Negative** tool.

Fig 3.3.9 Instructions set to play game



Instructions

Clicking on the **"Label All Negative"** button diagnoses **all remaining cells** as malaria-free. Only use this button when you think you have removed all infected and questionable cases.

When the game advances, either a new set of cells will appear on the screen for diagnosis or the game will end and you will be able to review your misdiagnoses.

Fig 3.3.10 Instruction set to play game

Fig 3.3.11 Main menu

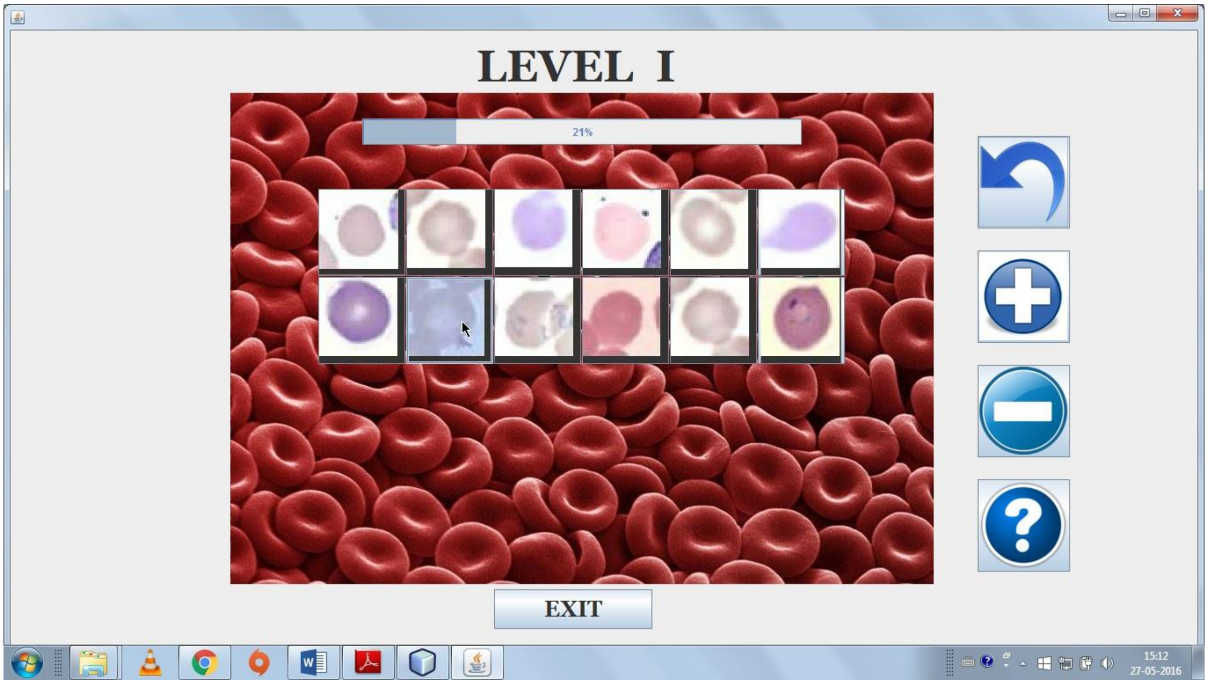


Fig 3.3.12 Gameplay

CHAPTER 4- PERFORMANCE ANALYSIS

4.1 Introduction

To test the credibility of our crowd-sourced gaming-based malaria diagnosis platform, we embedded RBC images as control images within the game such that each gamer had to go through all the RBC images.

Then we combine the diagnosis of all the gamers to measure accuracy, sensitivity and specificity of the responses generated by all the individuals.

The above discussed performance metrics are defined as a function of the effective crowd size, which we define as the minimum number of times that all of the cells in the dataset have been diagnosed by the gamers.

Sensitivity, Specificity, and Accuracy are the performance metrics terms which are generally associated with a binary classification test and they statistically measure the performance of the test involved. In a binary classification, we divide a given data set into two categories on the basis of whether they have a certain property or not by differentiating between the categories visually or by performing certain tests and in a binary classification test, as the name itself conveys, we deal with two datasets.

Of these two categories, in general, sensitivity indicates, how well the test predicts one category and specificity measures how well the test predicts the other category. Whereas accuracy is expected to measure how well the test predicts both categories.

A normal human being or the non-expert gamer in this case, has the potential to completely discriminate red blood cells with and without infection. Here, subjects are the images of red blood cells. Values of a perfect test which are above the cut-off are always indicating the disease, while the values below the cut-off are always excluding the disease. Unfortunately, such perfect test does not exist in real life and therefore diagnostic procedures can make only partial distinction between subjects with and without disease. Values above the cut-off are not always indicative of a disease since subjects without disease can also sometimes have elevated values. Such elevated values of certain parameter of interest are called false positive values (FP). On the other hand, values below the cut-off are mainly found in subjects without disease. However, some subjects with the disease can have them too. Those values are false negative values (FN). Therefore, the cut-off divides the population of examined subjects with and without disease in four subgroups considering parameter values of interest:

1. True positive (TP) – Number of correctly labelled positive samples.
2. False positive (FP) – Number of negative samples incorrectly labelled as positive.
3. True negative (TN) – Number of correctly labelled negative samples.
4. False negative (FN) – Number of positive samples incorrectly labelled as negative

Diagnostic accuracy measures tell us about the ability of a test to discriminate between and/or predict disease and health. It is calculated by (TP+TN/TP+TN+FP+FN). In general, measures of diagnostic accuracy are extremely sensitive to the design of the study.

Sensitivity is expressed in percentage and defines the proportion of true positive subjects with the disease in a total group of subjects with the disease (TP/TP+FN). Hence, it relates to the potential of a test to recognize subjects with the disease.

Specificity is a measure of a diagnostic test accuracy, complementary to sensitivity. It is defined as a proportion of subjects without the disease with negative test result in total of subjects without disease (TN/TN+FP). It describes the test ability to recognize subjects without the disease, i.e. to exclude the condition of interest.

4.2 Results

Our focus was mainly on calculating accuracy, sensitivity and specificity of the online gamers in estimating the image sets.

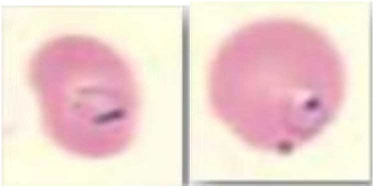
Since we are combining decisions that are received from many gamers, we will be feeding them the same data set of images to label. Therefore, there is a single sequence of images to be labelled, and each gamer will output a decision sequence. Ideally, the output of each gamer yields the correct diagnostic labels for the blood cell images. Given that each image either corresponds to a healthy cell or an infected cell, we can use binary labels to identify them: 0 for healthy and 1 for infected.

We should be clear about the part that throughout the project we discuss diagnosis results for 'individual' RBCs only. In reality, malaria diagnosis using a blood smear sample corresponding to a particular patient is a relatively easier task compared to single cell diagnosis since a thin blood smear for each patient's sample contains thousands of RBCs on it.

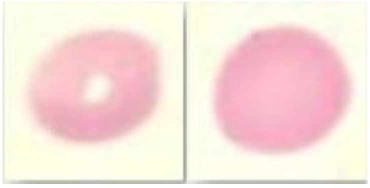
The specificity is always very high for the gamers, and does not improve much as more gamers are added to the mix. However, the sensitivity benefits the most as more players are added. The accuracy also increases as more players are added, but since it reflects both the specificity and the sensitivity its increase is not as drastic as that of the sensitivity.

Here TP=20, FP=1, FN=1, and TN=50. On putting these values in the formulae mentioned in previous section, we can calculate accuracy, sensitivity, and specificity.

The combined accuracy of the gamer diagnoses for 72 red blood cells images was 97.2%, with sensitivity (SE) of 95.24% and specificity (SP) of 98.04%.



Positive Samples



Negative Samples

CHAPTER 5- CONCLUSION

5.1 Discussion

In this project, we have utilized concepts of human computation and crowd sourcing to develop a revolutionary new framework for tele pathology, and creating a new paradigm for remote medical image diagnosis. This platform relies on the participation of non-expert crowds of individuals by playing bio-games that are designed around specific pathology tasks. To demonstrate the capabilities of this platform we have designed a game for distributed diagnosis of malaria.

We have shown that by utilizing the innate visual recognition and learning capabilities of human crowds it is possible to conduct reliable microscopic analysis of biomedical samples and make diagnostics decisions based on crowd-sourcing of microscopic data through intelligently designed and entertaining games that are interfaced with artificial learning and processing back-ends.

We demonstrated that in the case of binary diagnostics decisions (e.g., infected vs. uninfected), using crowd-sourced games it is possible to approach the accuracy of medical experts in making such diagnosis.

We have described how the task of binary image-based diagnosis can be crowd-sourced to minimally trained individuals and yet yield accurate results.

5.2 Future Scope

The work presented in this report is a proof of concept and not the complete envisioned system, with potentially thousands of gamers and many patient slides to be diagnosed, which is left as future work. In addition to generating remote biomedical diagnosis through engaging games, the Distributed Medical Diagnosis through Gaming presented platform can serve as an information hub for the global healthcare community.

This digital hub will allow for the creation of very large databases of microscopic images that can be used for e.g., the purposes of training and fine-tuning automated computer vision algorithms. It can also serve as an analysis tool for health-care policy makers toward e.g., better management and/or prevention of pandemics.

Next, we would like to briefly discuss regulatory and practical issues that need to be addressed for deployment of the presented gaming and crowd-sourcing-based diagnosis and telemedicine platform. As a potential future expansion of the platform, incentives can be used to recruit healthcare professionals who are trained and educated to diagnose such biomedical conditions, making them part of our gamer crowd.

In such a scenario, one can envision the gaming platform to serve as an intelligent telemedicine backbone that helps the sharing of medical resources through e.g., remote diagnosis and centralized data collection/processing.

In other words, it would be a platform whereby the diagnosis can take place by professionals far away from the point-of-care. At the same time, it also enables the resolution of possible conflicting diagnostics decisions among medical experts, potentially improving the diagnostics outcome.

5.3 Applications

We hope that this game will get widely used and expand its database as the Bio Games platform continues to create gold labels for new microscopic images of thin or thick blood smears. With large databases connected to user-friendly games and web-interfaces, this platform could be used for better training and education of medical personnel toward accurate reading of microscopic slides as well as for training of machine learning algorithms to automate digital diagnosis.

This approach can also be scaled to other diseases besides malaria. Considering the relatively poor training of health-care workers in developing countries, this approach could be especially valuable for improving the accuracy of malaria diagnosis and measurement of parasitemia which are infected patients that are on treatment.

Assuming the expansion of this crowdsourced diagnostics platform and the generation of large image databases with correct diagnostics labels, software can be created to make use of such databases to assist in the training of medical professionals. Through such software, medical students and/or trainees can spend time looking at thousands of images, attempting diagnosis, and getting real-time feedback on their performances. Based on the concepts described in this report, we also envision this platform to expand to other micro-analysis and diagnostics needs where biomedical images need to be examined by experts.

In conclusion, here we described how the task of binary image-based diagnosis can be crowd-sourced to minimally trained individuals and yet yield accurate results. We reckon that under many circumstances this methodology is more practical than the use of automated computer algorithms for the same purpose since the human visual system offers a very low-cost and highly superior pattern recognition platform.

Open problems that could be solved using human computing and crowd sourcing include [4]:

- **Language Translation**. Imagine a game in which two players that do not speak the same language work together to translate text from one language to the other.

- **Monitoring of Security Cameras**. With cameras becoming less expensive over time, it is now feasible to have security cameras everywhere. Imagine a game in which people watch the security cameras and alert authorities of illegal activity.

- **Improving Web Search**. Different people have different levels of skill at searching for information on the Web. Imagine a game in which the players perform searches for other people.

- **Text Summarization**. Imagine a game in which people summarize important documents for the rest of the world.

REFERENCES

[1] Crowd-sourced BioGames: managing the big data problem for next generation lab-on-a-chip platforms, Sam Mavandadi, Received 28th May 2012, Accepted 8th August 2012.

[2] Human Computation by Luis von Ahn CMU-CS-05-193December 7, 2005S.

[3] BioGames: A Platform for Crowd-Sourced Biomedical Image Analysis and Telediagnosis Sam Mavandadi, PhD, 1, 2 Steve Feng, BS, 1, 2 Frank Yu, BS, 1, 2 Stoyan Dimitrov, BS, 1, 2 Richard Yu, BS,1,2 and Aydogan Ozcan, PhD1–3.

[4] Harris, Christopher Glenn. "Applying human computation methods to information science." PhD (Doctor of Philosophy) thesis, University of Iowa, 2013.

[5] Distributed Medical Image Analysis and Diagnosis through Crowd-Sourced Games: A Malaria Case Study Sam Mavandadi, Stoyan Dimitrov, Steve Feng, Frank Yu, Uzair Sikora, Oguzhan Yaglidere, Swati Padmanabhan, Karin Nielsen, Aydogan Ozcan.

[6] von Ahn L, Dabbish L (2008) Designing games with a purpose. Communications of the ACM, 58–67.

[7] Tuite K, Snavely N, Hsiao Dy, Tabing N, Popovic Z (2011) Photo City: training experts at large-scale image acquisition through a competitive game. In: Proceedings of the 2011 annual conference on Human factors in computing systems, Vancouver, BC, Canada, 1383–1392.

[8] Tek FB, Dempster AG, Kale I (2010) Parasite detection and identification for automated thin blood films malaria diagnosis. Computer Vision and Image Understanding 114: 21–32.

[9] Optimization Techniques for Human Computation-enabled Data Processing By Adam Marcus, S.M., Massachusetts Institute of Technology (2008).