A Proposal for Automatic Diagnosis of Malaria

[Extended Abstract]

Allisson D. Oliveira
University Federal Rural of Pernambuco (UFRPE)
two Brothers, 52171-900
Recife-PE, Brazil
allissondantas@gmail.com

Giordano Cabral
University Federal Rural of Pernambuco (UFRPE)
two Brothers, 52171-900
Recife-PE, Brazil
giordanorec@gmail.com

D. López
Universitat Politécnica de Catalunya (UPC)
despatx 2508860
Barcelona, Spain
daniel.lopezcodina@upc.edu

Caetano. Firmo
University Federal Rural of Pernambuco (UFRPE)
two Brothers, 52171-900
Recife-PE, Brazil
caetanof rmo@gmail.com

F. Zarzuela Serrat
Institut Català de la Salut (ICS)
Av Drassanes 08011
Barcelona, Spain
francesc.bcn.ics@gencat.cat

J. Albuquerque
University Federal Rural of Pernambuco (UFRPE)
two Brothers, 52171-900
Recife-PE, Brazil
joa@deinfo.ufrpe.br

ABSTRACT

This paper presents a methodology for automatic diagnosis of malaria using computer vision techniques combined with artificial intelligence. We had obtained an accuracy rate of 74% in the detection system.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; D.2.8 [**Software Engineering**]: Metrics—*complexity measures, performance measures*; I.2.1 [**Artificial Intelligence**]: Miscellaneous

General Terms

Experimentation, Algorithms.

Keywords

Malaria, automatic diagnostic, detecting malaria, artificial intelligence, haar, computer vision techniques.

1. INTRODUCTION

Malaria is a worldwide public health problem. In 2010, according to the World Health Organization (WHO) approximately 24 million cases were diagnosed with positive incidence of the disease. A total of 106 endemic areas were identified and Africa was the most affected continent [6]. It is caused by a protozoan of the genus Plasmodium, where the vector is the female of the mosquito Anopheles. There are four types of Plasmodium that cause malaria in humans: P.vivax, P.falciparum, P.malariae and P.ovale. The species P. falciparum and P. vivax are responsible for 95% of cases reported worldwide [3].

Malaria is very strongly related to remote areas such as in Africa, a system of automatic diagnosis and low cost had become a basic tool if we really want to control it.

WWW 2013 Companion, May 13–17, 2013, Rio de Janeiro, Brazil.
ACM 978-1-4503-2038-2/13/05.

2. RELATED WORK

Crowd-sourcing approach has been used to pattern recognition which in essence seeks to solve complex computational problems. In this approach, the same problem is distributed to several participants who contribute to the resolution of the problem. Several gaming platforms have been developed to solve the problems as crowd-sourcing in biology and medical sciences. Thus this approach allows non specialists to solve such problems. A recent work for recognition of infected cells by malaria uses crowd-sourcing [5]. It is a game for recognition of cells infected with malaria, where the slides are given to non-specialists in malaria and in the background a learning system where artificial is made several comparisons with issues of human and machine learning. This work reports impressive results of the non-specialist players who reported malaria infection with an accuracy of 1.25% when compared to a specialist doctor.

3. METHODOLOGY

The facial recognition method proposed by Viola and Jones[7] is a known worldwide optimal algorithm for this purpose. However other studies like Firmo[1] demonstrate that it is possible to use this technique for disease detection in laboratory slides.

The method of Viola and Jones[7] uses rectangles as features to locate faces. These are rectangular objects with two light and dark regions. The calculation is given by the difference between the sum of the intensities of the pixels of clear region and the sum of the intensities of the pixels in the dark region $\sum \mathbf{R}_{clear} - \sum \mathbf{R}_{dark}$.

Rectangles are illustrated in Figure 1.

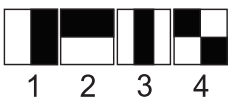

Figure 1: basic characteristics of haar

Rectangles can be processed as an Integral Image that is the representation of an image in your position (x, y). The

position contains the intensity value of all pixels above and to the left of point, including it its own location (x, y). For such is given in equation 1.

$$ii(x, y) = \sum_{x^l \leq x, y^l \leq y} i(x^l, y^l), \quad (1)$$

The evaluation of any feature rectangle can be made with 4 reference to the image. The sum of the pixels of the rectangles of arbitrary size can be calculated in constant time.

To extract the features most effective algorithm is used Adaboost, one of the features this algorithm is weight distribution on sets of examples and a modification this distribution during the course iterations of the algorithm, this algorithm is based in bosting, where strong hypotheses are formed by means of a linear combination of weak hypotheses. This proposes then a cascade of classifiers[1][7].

For the diagnosis of malaria were considered some steps that are respectively image acquisition,preprocessing image, training and classification.

In this first phase of the project image acquisition, we use a set of 13 images of malaria slides extracted by the laboratory group research *MOSIMBIO* of University Politecnica de Catalunya, identified 112 positive cases of malaria in this set of 13 images, using the method for preparing blood smear slides and a camera coupled to a microscope.

In the preprocessing image, in first step is used the image is converted in grayscale justified by the fact minimize the amount of calculation per pixel and noise reduction.In the second step histogram equalization that is a technique to distribute the grayscale values of pixels in an image, so as to obtain a uniform image[4][2].

The training is given a set of positive images containing face and a set of negative images that are not faces[7], in the case for malaria sets are defined infected with malaria (positive) and not infected with malaria (negative) as Figure 2 illustrates.

Figure 2: (a)positive sample grayscale(b)negative sample grayscale

The capture of image of training is overseen per human,that save objects of interest , a total of 147 sample positive and 520 negative sample was used to create the set of examples. This samples are partitioned into set of training and validation(76%) and test set(24%), the classifiers are formed using the conventional *haartraining process*[7][2].

4. RESULTS

The training detection system consists of sequenced classifiers based on the methodology proposed by Viola and Jones[7] in order to form a single strong classifier with the purpose of identifying malaria infection.

In his best performance reached It had obtained an accuracy rate of 74% in classification of true positives in classifiers with 10 stages, the set of tests was composed of 35 positive examples. We also carried out experiments with 10, 15, 22 and 25 classifiers.

5. CONCLUSIONS

This approach enables to train a technician with an automated system and develop a malaria diagnosis on site and in real time. This is very important for remote areas. This is accomplished through a mobile device already informing the result and medicating the patient or even refer to a treatment center or hospital data would be stored. The results can be send to a central epidemiological once the device to get internet connection.

6. ACKNOWLEDGMENTS

This work was partially supported by FACEPE and UFRPE and DEINFO.

7. REFERENCES

[1] AC.Firmo. Classificação adaboost com treinamento por enxame de partículas para diagnóstico da esquistossomose mansônica no litoral de pernambuco. Master's thesis, Universidade de Pernambuco, Recife, 2010.

[2] G. Bradski and A. Kaehler. *Learning OpenCV: Computer Vision with the OpenCV Library*. O'Reilly Media, 2008.

[3] L. S. Garcia. Malaria. *Clinics in Laboratory Medicine*, 30(1):93 – 129, 2010. Emerging Pathogens.

[4] R. González and R. Woods. *Digital Image Processing*. Prentice Hall, 2008.

[5] S. Mavandadi, S. Dimitrov, S. Feng, F. Yu, U. Sikora, O. Yaglidere, S. Padmanabhan, K. Nielsen, and A. Ozcan. Distributed medical image analysis and diagnosis through crowd-sourced games: A malaria case study. *PLoS ONE*, 7(5):e37245, 05 2012.

[6] W. H. Organization. Malaria, 2012.

[7] P. Viola and M. J. Jones. Robust real-time face detection. *Int. J. Comput. Vision*, 57(2):137–154, May 2004.