

A Neural Network Architecture for Automated Recognition of Intracellular Malaria Parasites in Stained Blood Films

Sobath Pradeepa Premaratne^a, Nadira Dharshani Karunaweera^b, Shyam Fernando^c,
W Supun R Perera^b, R P Asanga S Rajapaksha^b

^a Department of Surgery, ^b Department of Parasitology, ^c Department of Clinical Medicine,
Faculty of Medicine, University of Colombo, Sri Lanka.

Abstract

The global burden of malaria is enormous and the development of better laboratory diagnostic tools is a key step in malaria control recommended by the WHO. Our objective was to develop an automated tool for the recognition of intracellular malaria parasites in stained blood films.

We used digital images of oil immersion views from microscopic slides captured through a capture card. They were preprocessed by segmentation and grayscale conversion to reduce their dimensionality and later fed into a feed forward backpropagation neural network (NN) for training it. Then a user interface was developed incorporating this trained NN.

In the final product, the tool allows a user to view the slide in a graphical user interface. When the user gives a command to analyze, a still image is captured and sent to the NN for recognition after preprocessing.

Preliminary results show that the NN can identify carefully selected test data.

Key Words:

Neural networks, Malaria, Automated microscopy, Backpropagation image classification

Introduction

Malaria is a public health problem in more than 90 countries, inhabited by a total of some 2.4 billion people, representing about 40% of the world's population [1]. Best estimates currently describe the annual global burden of malaria as: 1.1 million deaths, 300-500 million cases and 44 million disability adjusted life years (DALYs)[2].

It has been estimated that the economic burden is also extremely high, accounting for a reduction of 1.3% in the annual economic growth rate of malaria endemic countries, and that the long-term impact in these countries is a reduction of GNP of more than half [3].

Malaria is caused by the single-celled protozoan parasites of the genus *Plasmodium*. Four species (*P. falciparum*, *P. vivax*, *P. ovale*, *P. malariae*) infect humans by entering the bloodstream [4].

Major trends over the last few decades point to a worsening situation if effective action is not taken. In 1999 WHO has put forward a control strategy for malaria, which focuses on early detection of cases, development of new tools, strategies and methodologies, and improvement of existing tools through research and development [5].

Laboratory Diagnosis of Malaria

Microscopy of Giemsa-stained thick and thin films remains the current standard for diagnosis of active malaria. Out of other laboratory techniques available, molecular diagnostics which employ PCR is expensive and malaria antibody detection can detect past infections but not necessarily active cases. Immunologic and biochemical detection products are still not cost effective and cannot indicate the parasite load [6].

Although microscopy has good sensitivity and allows species identification and parasite counts it requires microscopical expertise and is a labour intensive repetitive task which is time consuming.

The significance of malaria as a global health problem and the importance of laboratory diagnosis in malaria control, lack of non-microscopic methods of diagnosis and the problems inherent to microscopy prompted us to develop a tool which can automate the process of parasite detection in stain blood films.

Objective

To develop an automated tool for the recognition of intracellular Malaria parasites in stained blood films.

Materials and Methods

Artificial Neural Networks (ANN's) have proven to be a promising paradigm for Intelligent Systems. Neural networks have been trained to perform complex functions in various fields of application including pattern recognition, identification, classification, speech, vision and control systems [7].

ANN's have the advantage of learning by example and the ability to generalize from their training data to other data. They are fault tolerant in the sense, they can produce correct outputs from noisy and incomplete data. ANN's are relatively inexpensive to build and train [8].

These features of ANN's prompted us to look for a ANN based solution for the malaria parasite detection tool.

The project was initially planned to be completed in 2 stages. In the first stage it was decided to develop a suitable ANN and train it with a data set to find out the feasibility of using ANN for this project. The second stage was devoted to developing a suitable user interface incorporating the ANN and to create a real-time image acquisition system from the microscope.

Stage 1:

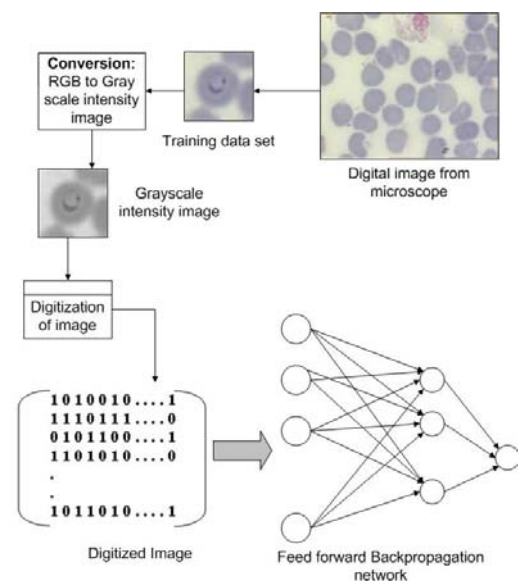
Oil immersion views (10x1000), of Giemsa stained blood films with the ring stage of the *Plasmodium falciparum* was captured using a binocular microscope mounted with a scientific video camera.

Only *Plasmodium falciparum* ring stages were chosen as almost all deaths from malaria are due to this species and the necessity to use only one stage of the lifecycle for the initial training of the ANN to keep things simple.

The analog signal from the video camera was converted to Digital Video using a capture card and compatible software. Captured images were 352 pixels X 255 pixels bitmap images.

Later they were segmented to 64 pixels X 64 pixels images to be used as a training data set. The size of 64 pixels X 64 pixels was chosen as it roughly corresponded to the size of a parasite infested red blood cell at the captured magnification. The other reason for the segmentation was to make sure that the ANN's was kept to the smallest possible size in order to achieve easier training.

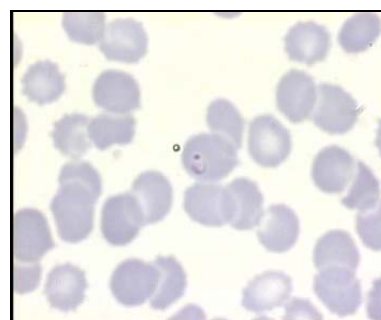
Thereafter images in the training data set were converted to Grayscale Intensity maps from their original RGB format, to reduce the dimensionality [9] of the input matrices to the ANN. Following this the images were digitized and the training of the ANN was done using these as input.



Feed Forward Backpropagation neural network architecture developed by Paul Werbos was chosen as it was a simple and one of the most commonly used ANN's [10].

In this type of ANN a new input could lead to correct output provided that the input being presented was similar to the inputs used in training the network. This generalization property makes it possible to train a network on a representative set of input/target pairs and get good results without training the network on all possible input/output pairs [10].

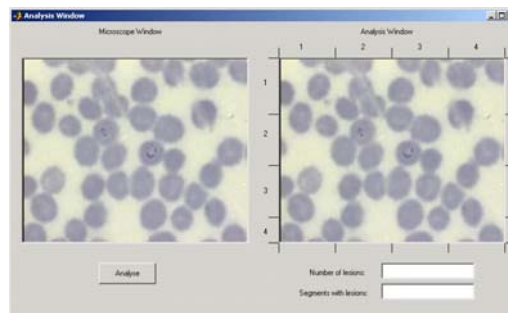
Another reason to chose backpropagation was it's ability to perform pattern classification on data where the input and the output had no linear relationship, as in the case of this application [11].



Thin blood film (x1000 - Giemsa)

Stage 2:

In the second stage a software tool was developed incorporating the ANN that was trained, which could also capture the video stream coming from the camera mounted on the microscope and display it to the user in a graphical user interface(GUI).



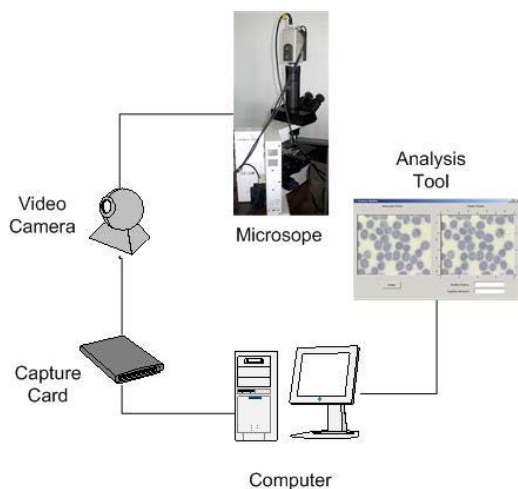
The same GUI could display the result on the analysis window after analyzing the captured image.

Final Product

In the final product, oil immersion view of the microscope was captured via a scientific video camera and converted to 'Digital Video' format using a capture card and displayed as a video clip to the user on the GUI.

The user can start the process of image analysis by pressing the analysis button on the GUI. This event triggers a series of secondary events.

First a single frame from the digital video stream is captured and sent for preprocessing where the image is segmented to 64 pixel X 64 pixel sub images. These are then converted to grayscale intensity images to reduce the dimensionality.



The images preprocessed by this method are then fed into the ANN that has been trained to recognize the malaria parasites. The ANN will categorize images to two categories, i.e. images with parasites and with out parasites. Image segments with malaria parasites are shown by highlighting them on the analysis window. The user is also presented with information such as how many image segments have parasites and their coordinates on the analysis window.

Results

The training data set for the ANN had 50 image segments with the malaria parasites and 50 images segments without. At the end of training the network achieved its performance function but the time taken to achieve it was significantly high.

When tested with a selected set of different images other than that used for training the ANN was able to categorize it accordingly. Since the training of the ANN is still in its early stages post training analysis has not been carried out.

When incorporated into the software tool the performance of the ANN was not satisfactory as there were a substantial number of errors in categorizing. This was expected and the speculated reasons are elaborated in the discussion.

Discussion

The training data set for the ANN was compiled with the help of human experts (experienced microscopists). In selecting a representative data set different intensities of staining and different shapes of the parasite's ring stage were taken into account.

Parasite positive image was defined as an image which had the full view of the parasite and hence images where part of the parasite was visible was taken as negative. This was done in order to increase the specificity of the tool.

For the training of the backpropagation ANN "Gradient Descent" algorithm with a performance function (error) goal of 10^{-3} and a learning rate of 0.05 was used in a batch processing mode.

The long time taken in training to achieve the performance function can be attributed to not using any optimization techniques such as the use of momentum technique and variable learning rates.

When used with the software tool in a practical scenario the results weren't satisfactory. This was expected as the ANN had not been trained with data directly from the tool.

Training with live data from the tool itself is the next goal and the use of different learning algorithms and learning rates with learning optimization techniques are yet to be undertaken.

This architecture of ANN based image analysis have numerous other applications in the field of medicine where humans experts are required to identify structures in microscopic slides (eg: microbiology, pathology).

Future Enhancements

As future enhancements it is planned to train the ANN to recognize different stages of the *P. falciparum* and also to train to recognize other species of *Plasmodium*. It is also planned to test with different architectures of ANN's such as "Radial Biasis".

Acknowledgements

We would like to acknowledge the support extended by the staff at the *Department of Parasitology* and *Department of Surgery*. A special mentioning is also due to *Mrs. H K Wickramathanthri* for her help in preparation of blood films.

References:

1. WHO fact sheet No.94, revised October 1998, World Health Organization, Geneva.
2. World Health Report 2002, World Health Organization, Geneva.
3. Sachs J. & Malaney P. (2002) The economic and social burden of malaria, *Nature* (415), 7: 680-685
4. Strategic direction for research. Tropical Disease Research. [2003 Mar] Available from: URL: <http://www.who.int/tdr/publications/publications/directi on.htm>
5. Malaria diagnosis - new perspectives, Report of a joint WHO/USAID informal consultation, October 1999
6. Malaria: Laboratory Diagnosis (Dec 2002) Available from: URL: <http://www.rph.wa.gov.au/labs/haem/malaria/diagnosis.html>

7. Jeffrey Johnson, Philipa Picton. Concepts in Artificial Intelligence; Butterworth-Heinemann LTD; 1995. Ch(4):95-96 ISBN:0-7506-2403-5;
8. Donna L Hudson, Maurice E Cohen; Neural Networks & Artificial Intelligence for Biomedical Engineering, IEEE Press; 200. ISBN: 81-203-1842-0
9. User guide, Image Processing Toolbox – For use with MATLAB. [Monograph available on CD Rom]. The MathWorks, Inc. Version 2.2; 1999
10. User guide, Neural Network Toolbox – For use with MATLAB [Monograph available on CD Rom]. Howard Demuth, Mark Beale. The MathWorks, Inc. Version 3.0; 1998
11. Valluru B Rao, Hayagriva V Rao; C ++ Neural Networks & Fuzzy Logic, 2nd Edition, BPB Publications; 1996. ISBN: 81-7029-694-3
12. Extracting Patterns of Lymphocyte Fluorescence from Digital Microscope Images, T. W. Nattkemper and H. Ritter and W. Schubert, Intelligent Data Analysis in Medicine and Pharmacology (IDAMAP 99) Workshop Notes , pp. 79-88 , addr. AMIA 1999 Washington DC , 1999

Correspondence

Dr. Sobath Premaratne,
No 459, Thimbirigasyay Road,
Colombo 5,
Sri Lanka.
sobath@lanka.ccom.lk