

A web-based intelligence platform for diagnosis of malaria in thick blood smear images: a case for a developing country

Rose Nakasi¹, Jeremy Francis Tusubira¹, Aminah Zawedde¹, Ali Mansourian², Ernest Mwebaze¹

¹Makerere University, Kampala, Uganda

²Lund University, Lund, Sweden

g.nakasirose@gmail.com, tusubirafrancisjeremy@gmail.com, sazawedde@gmail.com
ali.mansourian@nateko.lu.se, emwebaze@gmail.com

Abstract

Malaria is a public health problem which affects developing countries world-wide. Inadequate skilled lab technicians in remote areas of developing countries result in untimely diagnosis of malaria parasites making it hard for effective control of the disease in highly endemic areas. The development of remote systems that can provide fast, accurate and timely diagnosis is thus a necessary innovation. With availability of internet, mobile phones and computers, rapid dissemination and timely reporting of medical image analytics is possible. This study aimed at developing and implementing an automated web-based Malaria diagnostic system for thick blood smear images under light microscopy to identify parasites. We implement an image processing algorithm based on a pre-trained model of Faster Convolutional Neural Network (Faster R-CNN) and then integrate it with web-based technology to allow easy and convenient online identification of parasites by medical practitioners. Experiments carried out on the online system with test images showed that the system could identify pathogens with a mean average precision of 0.9306. The system holds the potential to improve the efficiency and accuracy in malaria diagnosis, especially in remote areas of developing countries that lack adequate skilled labor.

1. Introduction

Malaria is one of the leading causes of fatalities in the developing world [28]. This has been attributed to among others weak and poor diagnostic mechanisms and tools condemned for being slow, inaccurate and sinfully expensive for most developing countries [16]. Effective malaria diagnosis contributes towards effective disease management, control and surveillance. High-quality diagnosis is essential in all settings as misdiagnosis can result in drug resistance coupled with the economic burden of buying unnecessary

drugs and in some cases significant morbidity and mortality [16].

The standard system for diagnosis of Malaria plasmodium is microscopy of blood smear slides [3] [26]. Unlike Rapid Diagnostic Tests(RDTs) that detect specific antigens derived from malaria parasites in red blood cells [3], microscopy supports direct parasite detection [5]. The thick blood smear microscopy method is the most common method for screening malaria parasites. It facilitates the separating of plasmodium parasites from the Red blood cells associated with parasites in a sample of blood [24]. To achieve a negative or positive test, thick blood smear slides are usually used while the thin smears are used for species determination of the malaria infection. Laboratory diagnosis of Malaria mainly uses this method because of its simplicity, relatively low cost and the possibility of determining that magnitude of the parasitemia in a sample [24].

Microscopy, is not only time consuming (30 to 60 minutes) but requires skilled personnel to interpret results [11]. This makes conventional microscopy a subjective approach whose results vary significantly, a condition which is also blamed for low throughput screening especially in areas where there are few microscopists [6, 15]. A nationwide study in Ghana for example found 1.72 microscopes per 100,000 people and only 0.85 trained laboratory staff per 100,000 people [2] which is grossly inadequate. As a result, diagnoses are often made on the basis of clinical signs and symptoms alone. Diagnosis based on clinical signs and symptoms is not without challenges the most common being error-prone. The World Health Organisation alludes to this concern when it asserts that health problems may be compounded by variations in expertise on the one hand and health facilities resources such as laboratory equipment and test kits required to reliably diagnose diseases on the other hand [27].

These traditional microscopy related disadvantages have led to a pressing need for quality microscopy improve-

ments. Automation eliminates human error such as missed parasites in low parasitemia samples or fatigue that can occur during repeated viewing of stained slides. The emergence of web technologies and computer vision techniques that surpass human ability in detection of disease pathogens have the potential to improve public health. These can be used to bridge the diagnostic gaps in Malaria diagnosis. Web-based technology is used as a platform to enable easy and quick dissemination and interpretation of diagnostic information by users (lab technologists) appropriately. The integration of the web-based system and computer vision detection model will benefit laboratory technicians who wish to achieve decision-supported detection results. This mode of use can help laboratory staff to achieve consistency in diagnosis, and by focusing concentration on parts of the images likely to contain pathogens, may also help to relieve operator fatigue and improve throughput rates.

In this paper we present an efficient web-based intelligent malaria parasite detection system. A web-based system was integrated with a computer vision algorithm based on a pre-trained Faster R-CNN model to enhance diagnosis. Training and validation of the detection model was implemented on Google cloud infrastructure. The web-based application was implemented based on the Tensorflow object detection API [10] and used for testing. Automated web-based disease diagnosis infrastructure in this study aimed at aiding health communities especially in highly endemic and remote areas to efficiently diagnose malaria in a fast, accurate and timely manner.

2. Related work

The concept of web-based expert systems for malaria diagnosis has been in different aspects researched about. One such a system, the Medical Diagnosis Expert System (MDES) was implemented using the C Language Integrated Production System (CLIPS) [8]. CLIPS is an expert system, which has a shell composed of four modules: the user interface, the explanation system, the inference engine and the knowledge base editor. A similar malaria expert system prototype that involved a knowledge component, the application component (AC), the database system component (DC), the Graphical User Interface (GUI) component and the user component (UC) is also found in the literature [14]. For this system, the user interface component was implemented using the Java Programming language. The application component was implemented using the Java Expert System Shell (JESS) and the Java IDE built on Netbeans while the database component was implemented using SQL Server. However, in all these studies, the diagnosis engines are based on manual results from conventional lab tests.

Most of the previous attempts on computer vision based diagnosis of Malaria have been reviewed in numerous

scholarly literature for example [22]. Studies that have implemented Deep learning models [4, 17] have had promising results. The adoption of new state-of-the-art pre-trained deep learning models especially on thin blood smear images has also been studied in [19, 20, 18, 9].

Although the reported performance is promising, most of the existing research in transfer learning using deep learning models for malaria diagnosis has been only evaluated on classification tasks [12] of only thin blood smear images. Thin blood smear images tend to be associated with missing malaria parasites due to low parasitemia [22]. Moreover, a lot of the previous work does not go further to implement a web based system for efficient remote disease analysis. The work being presented in this paper is based on the automated identification and detection of malaria parasites in thick blood smears. A malaria parasite model was trained, and a corresponding web-based diagnostics system for malaria parasites in thick blood smear images was developed with the Faster Regional Convolutional Neural Network (Faster R-CNN) model as the decision support system.

3. Faster R-CNN for object detection

A Faster Regional Convolutional Neural Network (Faster R-CNN) is a network for object detection that takes in an image and outputs bounding boxes around objects of interest with class labels [21]. It contains a Region Proposal Network (RPN) that extracts the target boxes. Faster R-CNN is an improvement on the related Fast R-CNN in terms of speed and performance. This is because of its ability to extract target boxes and to detect the class of the obtained region. The advantage of using Faster RCNN Tensorflow model is that the training interval which it takes in order to create a checkpoint file along with acceptable loss values is decreased by half the time as opposed to using architectures like Single Multibox Detector (SSD). Additionally, it also detects a higher amount of objects per image as well as increased accuracy for detection. Figure 1 shows the architecture of the Faster R-CNN model showing the two main components, the Regional Proposal Network and the detection CNN.

4. Proposed procedure for automated diagnosis

Intelligent diagnostic systems allow for effective and real time diagnosis of malaria. The fundamental question this paper thus attempts to address is. How best can malaria be efficiently diagnosed remotely in the developing countries that do not have adequate Microscopists?

The prototype system presented here implements a web-based detection system for diagnosis malaria from thick blood smears. Using web technology and computer vision, the architectural design shown in Figure 3 creates schematic

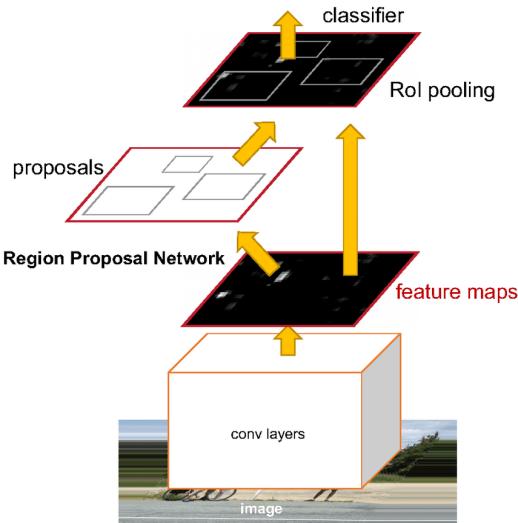


Figure 1. Faster R-CNN Architecture from [23] depicting the major blocks of the model architecture.

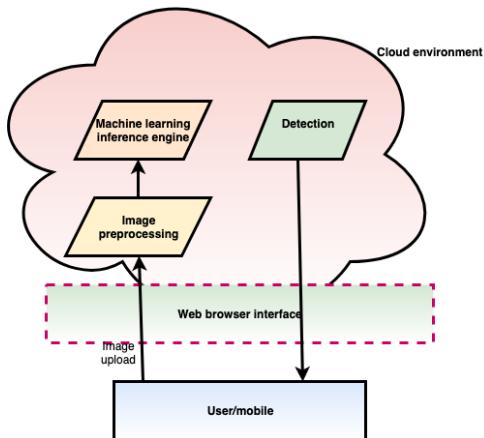


Figure 2. Architectural design for malaria automated diagnosis

for malaria diagnosis. We leverage the TensorFlow Object Detection API [10] to build the web-based malaria detection system.

5. Methodology

5.1. Data collection

Following the same Data collection procedure from previous work [17] on microscopic malaria diagnosis, 643 images of thick malaria blood smears were used. The images were of dimension 750 x 750 pixels and taken using a smartphone camera that was attached on the microscope eyepiece using a 3D printed adapter. In order to train the detection model, it was necessary to label the images and this was done with the help of expert lab technicians from Mulago Hospital in Uganda who manually drew bounding boxes

around the malaria parasites using an open source tool (LabelImg) [25].

The annotations obtained were saved in the PASCAL VOC format [13] with an image and its corresponding xml file. The corresponding xml file contained coordinates of bounding boxes labelled in the images. Every image contained at least one bounding box signifying presence of a parasite in the image. The annotated images were randomly split into a train and test set in the ratio of 9:1 respectively and the combined dataset was encoded in the tfrecord format that is optimized for processing with Tensorflow [1].

5.2. Data pre-processing

Experiments were performed on the dataset of thick blood smear images for malaria detection. Due to our limited dataset, data augmentation was applied. The primary purpose of the augmented training was to induce variation in the images, forcing the models not to get stuck in local minima and to increase generalisation as the models inherently are forced to learn a broader spectrum of spatial relationships.

Data augmentation was implemented by applying horizontal flipping and vertical flipping to randomly selected image samples. Subsequently, this form of augmentation strategy flips an image horizontally and vertically with a probability and comes embedded as helper functions in the pre-processing stage of the Tensorflow Object Detection API that was employed.

5.3. Model selection and training

Transfer learning was used to train the model. The advantage with transfer learning is that instead of starting the learning process from scratch, one starts from a point where the model has already learnt some patterns by training on other datasets to solve different problems. So basically, an already pre-trained model is selected, the output layer is then adjusted to match the number of labels of the new dataset.

In this study, the model trained on a dataset of 579 images and testing dataset of 64 images (90%:10%) respectively. Ubuntu system was used with a 5th Gen I7 Intel Core processor and 16GB RAM, a GPU-enabled Nvidia GTX 1060 Graphical Processing Unit (GPU) with 6GB RAM, Python 2.7 and a Tensorflow backend.

Training the models was done using open-source library Tensorflow [1] as it enabled convenient transfer learning from the base model to our specific models by loading the pre-trained models from the ImageNet dataset. Tensorflow also comes with an API to predict bounding boxes for detection of objects. The object detection models were downloaded from the Tensorflow Model Zoo [10], a holding repository for all current models in both object detection and image classification and pre-trained on a COCO dataset

[10]. The model was trained for up to 15,000 time steps until it converged.

5.4. Malaria Web based detection system

The application consists of two parts; the front-end and back-end. The front end was developed using Flask, a python framework for web development. It provided a browser interface which enables upload of an image and later returns an image with bounding boxes with confidence scores drawn around detected parasites.

The back-end which is essentially the brain of the App, hosts the TensorFlow model for predicting parasites, and a protobuf model file was served using Tensorflow Serving Engine. A Flask API was implemented on top of Tensorflow serving engine to provide a means of passing input and getting output from the model file (protobuf).

When an image is uploaded through the browser, it is converted into an array which is then sent via a Flask API to be processed by the model file. The output from the model file are coordinates of the parasites that have been detected on the input image along with the confidence score of each detection and these are sent via the Flask API back to the front-end. At the front-end, a function is implemented to draw the bounding boxes on the input image as highlighted in the architectural design (see Figure 3).

6. Results

6.1. Implementation of prototype system

Once an image is uploaded through a web browser, it is then sent for processing by the model and the output is a returned image with parasites detected characterised by bounding boxes with a level of accuracy confidence around the parasite. Each time a thick blood smear image is uploaded, the same logic is performed and if a parasite is detected, its bounding box is drawn highlighting the detection confidence. The final web based user interface is shown in Figures 4, 5 and 6.

6.2. Model performance evaluation

To ascertain the level of performance of the state-of-the-art pre-trained Faster R-CNN ResNet50 meta-architecture used in this study for malaria parasite detection in thick blood smear, the model was evaluated using mAP (mean Average Precision), precision and recall.

The resultant mAP@0.5, precision and recall performance shown in Table 1 indicates promising results from the two stage detector (Faster R-CNN) used in this study. This is majorly because two stage detectors are generally able to localise objects of interest resulting into a higher Intersection over Union (IoU) thus a higher mAP [7].

Performance metric	Accuracy
mAP@0.5	0.9306
Precision	0.7461
Recall	0.9062

Table 1. mAP@0.5, precision and recall performance for pre-trained Faster R-CNN ResNet 50 model used in the study.

7. Conclusion

In this study, a web based intelligent diagnosis system for malaria parasite detection was developed. A Faster R-CNN model with a ResNet50 backbone meta-architecture combined with web technology was used to create the intelligent inference engine that malaria parasite detection from thick blood smears. Model results showed that the diagnosis system performs quite well with a mean average precision above 90%. The prototype system provides a user-friendly interface that allows easy upload of test images and quick retrieval of diagnosis results thus depicting feasibility to remote field situations. The prototype system can be accessed at <http://18.218.239.12> and the implementation to recreate the experiments in the paper is available at <https://github.com/rnakasi/Web-based-Malaria-diagnosis>.

Acknowledgments

The authors thank Ministry of Health, Uganda for providing authorisation to use the data. They also acknowledge the collaborators from Mulago Referral Hospital, Uganda especially Dr. Alfred Andama and Vincent Wadda for granting us the medical and technical support in data acquisition, annotation and interpretability of results. In a special way, we appreciate the Swedish International Development Cooperation Agency for funding this work.

References

- [1] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, G. M. Derek, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, X. Zheng, and Y. Brain. Tensorflow: A system for large-scale machine learning. In *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, pages 265–283, 2016.
- [2] I Bates, V Bekoe, and A Asamoah-Adu. Improving the accuracy of malaria-related laboratory tests in ghana. *Malar. J.* Vol.3. No. 38, 2004.
- [3] C Beadle, G.W. Long, P.D. McElroy, S.L. Hoffman, W.R. Weiss, and A.J. Oloo. Diagnosis of malaria by detection of plasmodium falciparum hrp-2 antigen with a rapid dipstick antigen capture assay. *The Lancet*, 343(8897), 564–568, 1994.

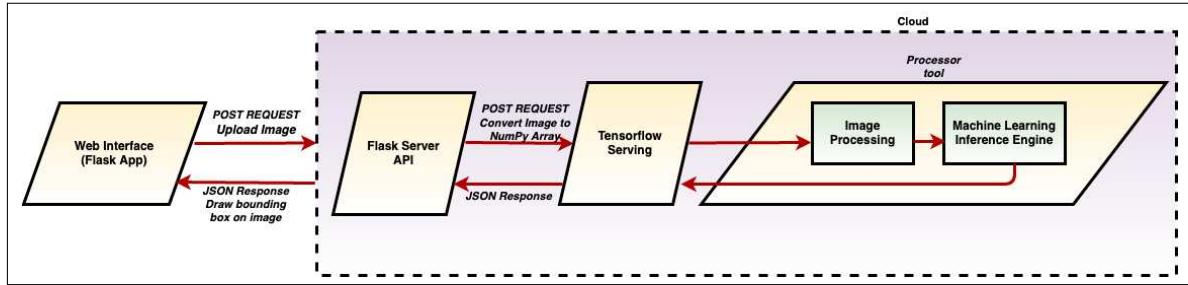


Figure 3. Architectural design of the malaria web-based detection system development.

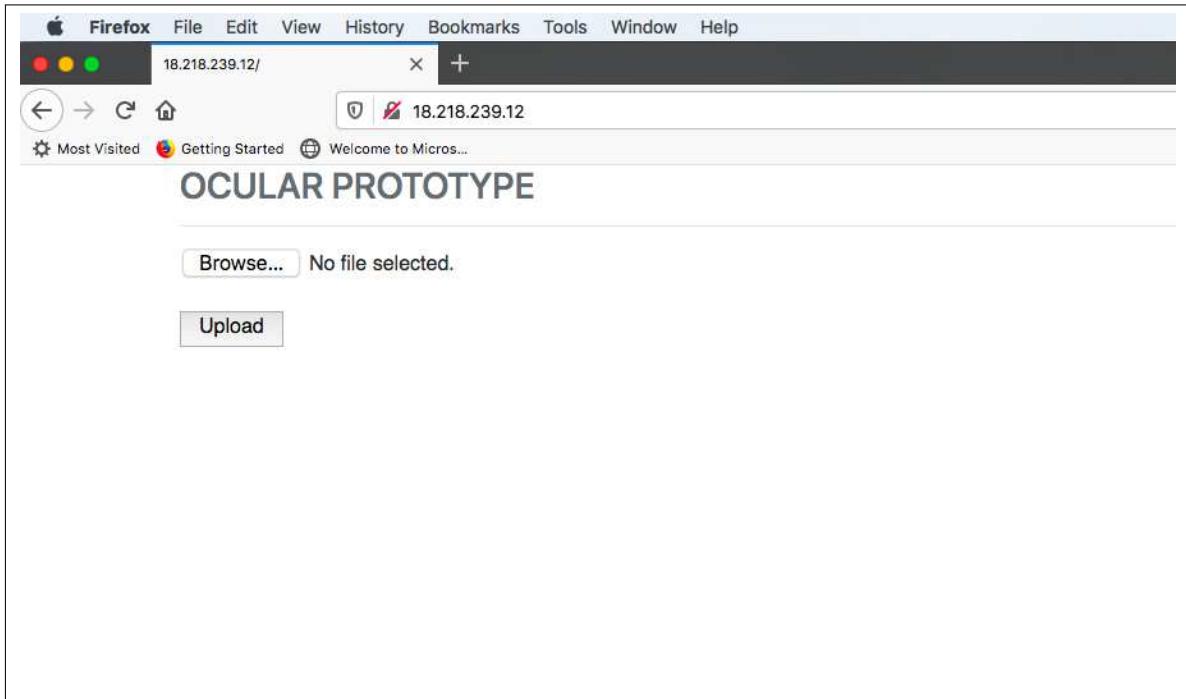


Figure 4. Browser interface that allows one to browse and upload microscopic images to the server.

- [4] P. Bengio, Y. and Vincent and A. Courville. Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 35(8): 1798-1828. *arXiv:1206.5538*. doi:10.1109/tpami., 2013.
- [5] D Bhattacharya. Relevance of economic field microscope in remote rural regions for concurrent observation of malaria and inflammation,. *Advances In Infectious Diseases (AID)*, Vol.2. No.1, PP. 13-18., 2012.
- [6] S Bowman and M.J Rhia. Impact of electronic health record systems on information integrity: quality and safety implications . 2013.
- [7] J. Cartucho. map. <https://github.com/Cartucho/mAP>. Accessed on 12 July 2019, 2019.
- [8] G. A. Chukwudebe, E. Ekwuuwune, and K. I. Nkuma-Udah. Medical diagnosis expert system for malaria and related diseases for developing countries. In *2017 IEEE 3rd International Conference on Electro-Technology for National Development (NIGERCON)*, pages 24–29, 2017.
- [9] Y. Dong, Z. Jiang, H. Shen, P. W. David, L. A. Williams, V.V.B. Reddy, W.H. Benjamin, and A.W. Bryan. Evaluations of deep convolutional neural networks for automatic identification of malaria infected cells. *IEEE EMBS international conference on biomedical and health informatics, BHI 2017*. Piscataway: IEEE, 101–104 DOI 10.1109/BHI.2017.7897215., 2017.
- [10] GoogleAI. Tensorflow object detection preprocessor. https://github.com/tensorflow/models/blob/master/research/object_detection/core/preprocessor.py L443. Accessed 17 June 2019, 2019.
- [11] M.M Kettelhut, H Edwards, A Moody, and P.L Chiodini. External quality assessment schemes raise standards: evidence from the ukneqas parasitology subschemes. *J Clin Pathol*. 56: 927-932. 10.1136/jcp.56.12.927., 2003.
- [12] M.M. Najafabadi, F. Villanustre, T.M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic. Deep learning applications and challenges in big data analytics. *Journal of Big Data*, DOI 10.1186/s40537-014-0007-7, 2015.

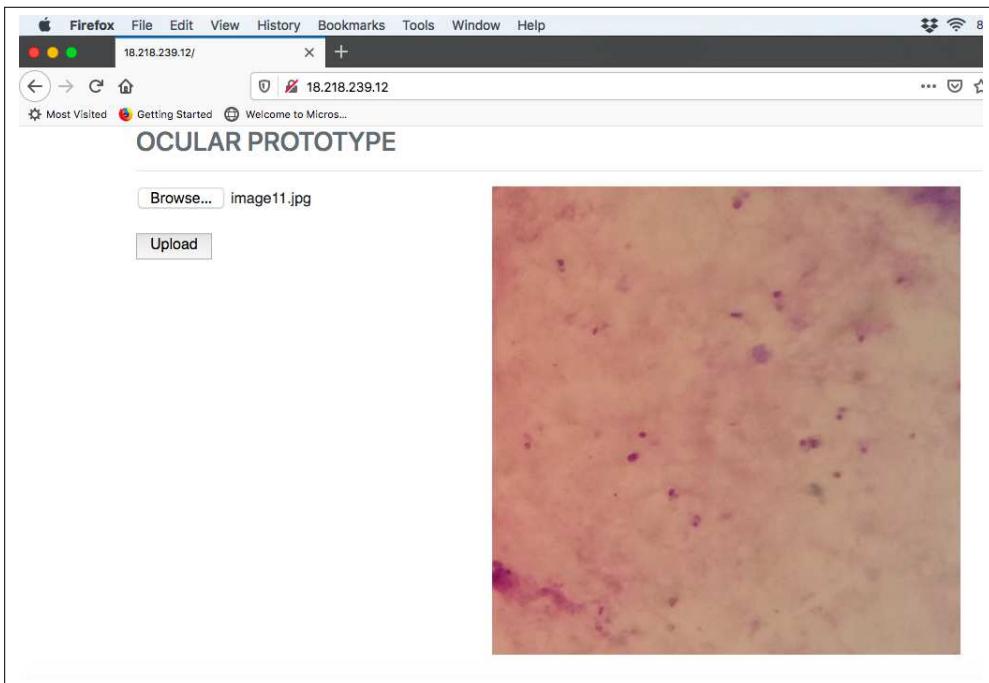


Figure 5. A microscopic thick blood smear image uploaded to the web browser.

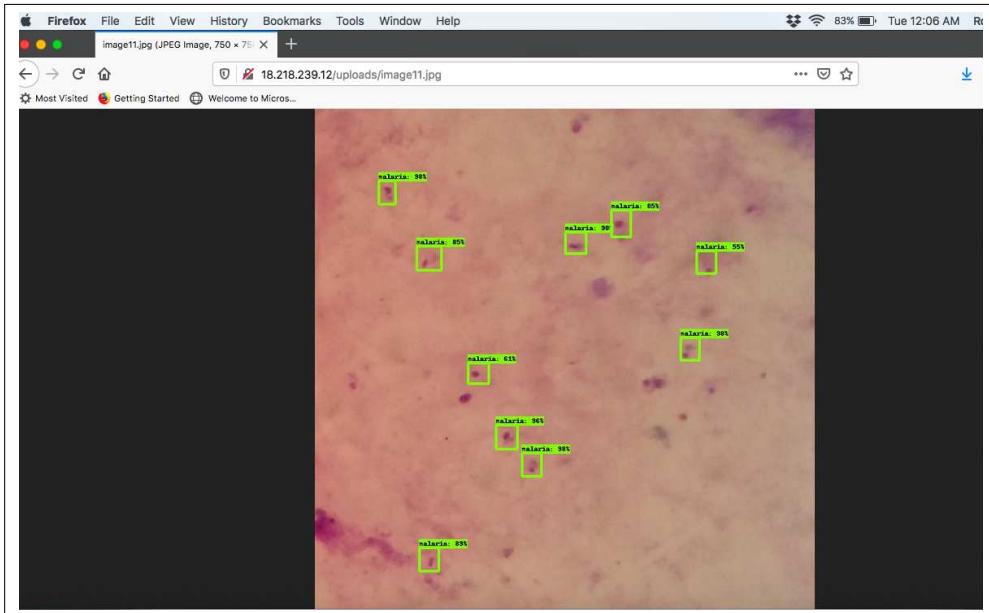


Figure 6. Automated detection of malaria parasites with confidence and bounding box in a microscopic image.

- [13] Pascal Network of Excellence. Pascal voc. <http://host.robots.ox.ac.uk/pascal/VOC/>. Accessed on 12 July 2019, 2019.
- [14] O Oluwagbemi, E Adeoye, and S Fatumo. Building a computer-based expert system for malaria environmental diagnosis: An alternative malaria control strategy. *Egyptian Journal of Computer Science*, 33, 10 2009.
- [15] A.P Patil, P.W Gething, F.B. Piel, and S.I. Hay. Bayesian geostatistics in health cartography: the perspective of malaria. *Trends Parasitol.* 2011, 27: 246-253. [10.1016/j.pt.2011.01.003](https://doi.org/10.1016/j.pt.2011.01.003), 2011.
- [16] C.A Petti, C.R Polage, T.C. Quinn, A.R Ronald, and M.A. Sande. Laboratory medicine in africa: A barrier to effective health care. *Clinical Infectious Diseases*, Vol. 42, NO. 3, PP.

377-382, 2006.

- [17] J.A. Quinn, R. Nakasi, P.K.B. Mugagga, P. Byanyima, W. Lubega, and A. Andama. Deep convolutional neural networks for microscopy-based point of care diagnosis. *In proceedings of International Conference on Machine Learning for Health Care, Volume 50.*, 2016.
- [18] S. Rajaraman, Sameer K., S.K. M. Antani, K. Silamut, M. A. Hossain, J. Richard, R.J. Maude, S. Jaeger, and G. R. Thomas. Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. *PeerJ.*; 6: e4568. ,doi: [10.7717/peerj.4568], 2018.
- [19] H Ravishankar, P Sudhakar, R Venkataramani, S Thiruvenkadam, Babu N Annangi, P, and V. Vaidya. Understanding the mechanisms of deep transfer learning for medical images. *GE Global Research, Bangalore, India.*, 2017.
- [20] A.S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson. Cnn features off-the-shelf: an astounding baseline for recognition. *IEEE computer society conference on computer vision and pattern recognition workshops; Piscataway.*, 2014.
- [21] S Ren, K He, R Girshick, and J Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 91–99. Curran Associates, Inc., 2015.
- [22] L. Rosado, L. Correia da Costa, J.M. Elias, and D. Cardoso. A review of automatic malaria parasites detection and segmentation in microscopic images. *Anti-Infective Agents*, 14(1):11–22., 2016.
- [23] Sik-Ho, T. Review: Faster r-cnn (object detection). https://miro.medium.com/max/1096/1*e6dx5qzUKWwasIVGSuCyDA.png, 2018. [Online; accessed Feb 25, 2020].
- [24] N Tangpukdee, C Duangdee, P Wilairatana, and S Krudsood. Malaria diagnosis: a brief review. *The Korean Journal of Parasitology*, 47(2):93–102, 2009.
- [25] D. Tarlin. Labelimg. <https://github.com/tzutalin/labelImg>., note = Online; Accessed 25 June 2019., year = 2019.
- [26] WHO. The role of laboratory diagnosis to support malaria disease management: Focus on the use of rapid diagnostic tests in the areas of high transmission. *WHO report*, 2004.
- [27] WHO. Malaria microscopy quality assurance manual. *WHO Press, World Health Organization. Geneva, Switzerland*, 2009.
- [28] WHO. World malaria report. *Geneva, Switerzerland, World Health Organisation.*, 2016.