

# 8

---

## *Automated Blood Smear Analysis for Mobile Malaria Diagnosis*

**John A. Quinn**

*College of Computing and Information Sciences  
Makerere University, Uganda*

**Alfred Andama**

*College of Health Sciences  
Makerere University, Uganda*

**Ian Munabi**

*College of Health Sciences  
Makerere University, Uganda*

**Fred N. Kiwanuka**

*College of Computing and Information Sciences  
Makerere University, Uganda*

### CONTENTS

8.1	Introduction .....	2
8.2	Conventional microscopic diagnosis of malaria .....	3
8.2.1	Practical difficulties in under-resourced health facilities .....	4
8.3	Alternative diagnosis methods .....	5
8.3.1	Rapid diagnostic tests .....	6
8.3.2	Related work in computer vision diagnostics .....	6
8.4	Blood smear image capture and annotation .....	7
8.4.1	Annotation .....	7
8.5	Automated diagnosis .....	8
8.5.1	Connected component features .....	10
8.5.2	Moment features .....	11
8.5.3	Classification .....	11
8.6	Evaluation .....	12
8.7	Discussion .....	14
8.8	Acknowledgements .....	15

## 8.1 Introduction

The gold standard test for malaria is the hundred-year-old method of preparing a blood smear on a glass slide, staining it, and examining it under a microscope to look for the parasite genus *plasmodium*. While several rapid diagnostic tests are also currently available, they still have shortcomings compared to microscopical analysis [18]. In the regions worst affected by malaria, reliable diagnoses are often difficult to obtain, and treatment is routinely prescribed based only on symptoms. Accurate diagnosis is clearly important, since false negatives can be fatal, and false positives lead to increased drug resistance, unnecessary economic burden, and possibly the failure to treat diseases with similar early symptoms such as meningitis or typhoid. The scale of the problem is huge: annually there are 300-500 million cases of acute malaria illness of which 1.1-2.7 million are fatal, most fatalities being among children under the age of five [27, 21, 22].

The lack of access to diagnosis in developing countries is largely due to a shortage of expertise, with a shortage of equipment being a secondary factor. For example, a recent survey carried out in Uganda [34] found 50% of rural health centres to have microscopes, but only 17% had laboratory technicians with the training necessary to use them for malaria diagnosis. Even where a microscopist is available, they are often oversubscribed and cannot spend long enough examining each sample to give a confident diagnosis.

This situation has prompted an increasing interest in finding technological solutions to carrying out the diagnosis automatically with computer vision methods, taking advantage of existing equipment and compensating for the shortage of human expertise. In particular, image processing and computer vision techniques can be used to identify parasites in blood smear images captured through a standard microscope. Given sufficient training data, the algorithms used in other medical imaging problems or computer vision tasks such as face detection can be applied to recognize plasmodia. Some studies have looked further at classifying the species and life cycle stage of parasites.

Apart from the idea of using blood smear images captured directly from a microscope, there is a great deal of attention currently on other forms of point of care diagnosis for malaria. Some of these are reviewed in the next section, and include methods based on fluorescence imaging or flow cytometry, for example. While these methods may be promising in future, there is still value in diagnosis based on image processing currently, for the following reasons:

- Image processing methods can be used when we do not wish to remove human experts from the diagnostic process completely, but rather to offer decision support. In this case, we might display to a technician (either on site, or remotely) the regions in blood smear images which seem most indicative of plasmodium, and allow the technician to make the final judgment. This could improve the efficacy of technicians by help-

ing to triage their attention, or make remote diagnosis over a network connection more feasible.

- When mobile devices are used for imaging and processing, we can take advantage of existing hardware. Both microscopes and camera phones are common in most malaria-affected countries. Hence, the only new hardware necessary to combine them is an adapter to mount the phone onto the microscope eyepiece or trinocular tube, which is relatively inexpensive. There have also been recent advances in low-cost microscopy using simple optical components attached to mobile devices [33, 6].
- Several other tests can be carried out with the same images, for example cell counts or detection of other hemoparasites. Malaria diagnosis might be just one element of a suite of diagnostic software running on the same system. In principle, any microscopic test could be automated with the same imaging hardware given sufficient training examples.

The rest of the chapter is organized as follows. In the next sections we review existing work on point of care diagnosis for malaria, and the standard practice for malaria diagnosis. We then describe a typical image capture setup, including experiments with 3D-printed phone adapters. Next, we describe methodology for extracting statistical image features, and application of learning algorithms to carry out malaria diagnosis as an object detection problem. Finally, we provide some quantification of the accuracy of the system, and conclude with a discussion of current issues and future directions.

---

## 8.2 Conventional microscopic diagnosis of malaria

The fundamental goal of malaria diagnosis is to demonstrate the presence of plasmodium before antimalarial drugs are used. Presumptive diagnosis from symptoms alone has poor accuracy and can lead to over-diagnosis of malaria, with resultant poor management of non-malarial febrile illnesses and wastage of antimalarial drugs [19]. Definitive diagnosis of malaria infection is still based on finding malaria parasites microscopically in stained blood films.

In thin films the red blood cells are fixed so the morphology of the parasitized cells can be seen. Species identification can be made, based on the size and shape of the various stages of the parasite and the presence of stippling (i.e. bright red dots) and fimbriation (i.e. ragged ends). However, malaria parasites may be missed on a thin blood film when there is a low parasitemia. Therefore, examination of a thick blood smear, which is 20–40 times more sensitive than thin smears for screening of plasmodium parasites, with a detection limit of 10–50 trophozoites/ $\mu\text{L}$  is recommended [28]. In this process, the red blood cells are lysed and diagnosis is based on the appearance of the

parasites which tend to be more compact and denser than in thin films. Since the thick smear is approximately 6-20 times as thick as a single layer of red blood cells, this results in a larger volume of blood being examined.

A group of dyes known as Romanowsky stains are a series of blue/red stains where the blue (methylene blue) binds to acidic substances and the red (eosin) binds to neutral or basic substances in cells. Examples of such stains include Fields A and B, Giemsa, Leishman and Wrights stain. Developed in 1800s by a Russian physician, these stains have similar basic components but differ from each other according to simple modifications.

While Leishman's stain (1901) undoubtedly gives the best results in a thin film, Giemsa stain (1902) has proved to be the best all-round stain for the routine diagnosis of malaria. It has the disadvantage of being relatively expensive, but this is outweighed by its stability over time and its consistent staining quality over a wide range of temperatures. The detection threshold in Giemsa-stained thick blood film has been estimated to be 4-20 parasites/ $\mu\text{L}$  [20]. Under field conditions, a threshold of about 50-100 parasites/ $\mu\text{L}$  blood is more realistic [16]. However, in remote settings with less skilled microscopists and poor equipment, a still higher threshold is likely.

The method preferred for staining thick blood smears in countries such as Uganda is Fields stain, particularly because it is more rapid than the alternatives. This stain is made of two components: Fields A is a buffered solution of azure dye and Fields B is a buffered solution of eosin. These stains are supplied ready to use by the manufacturer, and have advantages of being inexpensive, simple to use, economical and have short staining time compared to other methods. However there are also disadvantages with Field's stain, especially in under-resourced health centres in which the stain might be used. Poor blood film preparations often result in the generation of artifacts commonly mistaken for malaria parasites, such as bacteria, fungi, stain precipitation, and dirt and cell debris. These can frequently cause false positive readings [11].

### **8.2.1 Practical difficulties in under-resourced health facilities**

The procedure of staining with Field's A and B involves dipping the slide into the solution. This normally requires pouring the stains in couplin jars, which ideally should have lids, and water between them. These solutions are supposed to be prepared fresh each day for optimum potency, or kept tightly closed and filtered every morning. However, in resource constrained settings—under which the majority of health facilities in Uganda fall—these jars could be kept for weeks, and the stains rarely filtered. Frequent opening of these jars results in evaporation of methanol, which then results into precipitation artifacts. Contamination can also result from stain or water carry-over from one jar to the other (dilution). Bacterial contamination can originate from frequent opening of the jars or introduced from dirty slides, and if not filtered

regularly can result into false positive reporting as bacterial cells could be confused with malaria parasites.

Another contributing factor to poor microscopy performance is excessive workload, which is a problem when there is a shortage of staff with sufficient expertise. Sensitivity is directly related to the time available to examine blood films and therefore decreases when the number of slides exceeds the workload capacity of the microscopist, and this becomes more pronounced if the microscopist has responsibilities for diagnosing other diseases.

The time required to read an individual malaria slide depends on several factors, including the quality of microscope and immersion oil used, the skill of the microscopist, slide positivity rate (SPR) and parasite density. The time taken to declare a slide positive or negative differs considerably. A strongly positive thick film can be examined more quickly than a weak positive or more still a negative film. Another significant factor is the additional time required for species differentiation, where this is clinically important. Species identification is best done on a thin smear, but in low parasite density is extremely time-consuming.

During the malaria eradication era, the World Health Organization recommended that a single technician can satisfactorily read 60–75 slides per day. More recently, this has been considered unrealistic as a standard since the functions and roles of microscopists in malaria control are different today. Based on the Ugandan situation, the Malaria Control Programme together with Malaria Consortium have recommended 25-40 slides per day, each covering about 200 microscopic fields from a standard thick smear using x100 objective. This should take 5-10 minutes on average to declare a smear negative [4]. However, according to our personal experience, even after examining 20 consecutive slides, fatigue becomes an issue and concentration tends to weaken.

Perhaps the most important constraint for microscopy-based diagnosis in the developing world, however, is the frequent absence altogether of laboratory technicians from health facilities in rural areas. In Uganda, with only 34% of the laboratory staff positions occupied, the absence of this cadre of staff has been associated with higher costs of diagnostics based on microscopy in comparison with those based on RDTs [36, 5]. The absence of human resource for health is a major problem in sub-Saharan Africa, and one which gets worse in the lower health units which are commonly the first point of contact for patients.

---

### 8.3 Alternative diagnosis methods

The absence of human resource for performing the diagnosis of malaria in these setting is one of the reasons for the development of alternatives methods of di-

agnosis. This includes the recent development of several rapid diagnostic tests (RDTs) for malaria, methods for automating the microscopical examination with image processing, and other forms of diagnosis. In this section we review alternatives to the conventional diagnosis methods.

### 8.3.1 Rapid diagnostic tests

Rapid diagnostic tests (RDTs) for malaria have been a great success in reducing the disease burden following the change in policy by WHO [12, 35]. RDTs, based on testing for antigens produced by the immune system in response to plasmodium, have high sensitivity for parasite concentrations of over 500/ $\mu\text{L}$ . For smaller concentrations the sensitivity of RDTs becomes too low to be used reliably, however [36]. Test results are usually available in 5-20 minutes, do not require capital investment, electricity, or extensive training for laboratory staff, although individual tests are more expensive with RDTs than microscopical analysis [5].

Apart from the issues of inadequate sensitivity for low parasite concentrations, there are other concerns about the discriminative effectiveness of RDTs in specific situations. These include frequent false positive results in areas of low transmission [2] and false negatives for individuals with asymptomatic infections or multiple organism infestations [14]. Overall, RDTs are successful in a number of situations, but the gold standard for diagnosis for malaria remains the microscope—especially in those instances such as treatment failures or low parasitemia where RDTs will not work [1, 13].

### 8.3.2 Related work in computer vision diagnostics

A number of studies have looked at image processing and computer vision methodology for automated diagnosis of malaria from blood smears. In vision terms this is an object detection problem, and some previous work is reviewed in [31]. There has also been work in comparing these methods with other forms of diagnosis [3]. [24] uses neural networks with morphological features to identify red blood cells and possible parasites present on a microscopic slide. The results obtained with this technique were 85% recall and 81% precision using a set of 350 images containing 950 objects. In [30] a distance weighted  $k$ -nearest neighbor classifier was trained with features extracted by use of a Bayesian pixel classifier which was used to mark the stained pixels. The results achieved by this method were 74% recall and 88% precision.

Color space and morphological heuristics were employed to segment red blood cells and parasites by using an optimal saturation threshold [15] using a set of 55 images. Multi-class parasite identification, attempting to classify the type and life cycle stage of detected parasites has also been attempted [32].

All the vision systems mentioned operate on images of thin blood films, a single layer of blood cells with red blood cell matter preserved. Thick blood film images, which are prepared in such a way that red blood cell matter is

destroyed and DNA material is stained, are more commonly used in the developing world as they are more sensitive. This makes diagnosis possible even with very low parasitaemia, although typing of parasites is difficult with this type of sample as shape information is not as well preserved. Furthermore, the samples used in these previous studies were prepared under ideal conditions, with high quality slide preparation and imaging equipment. In the experiments described here, we use thick film images collected under field conditions.

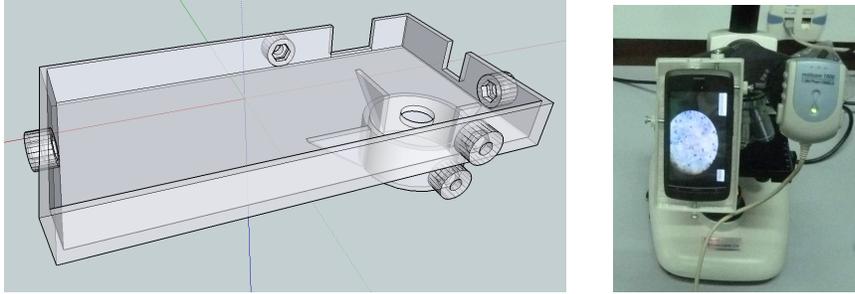
---

## 8.4 Blood smear image capture and annotation

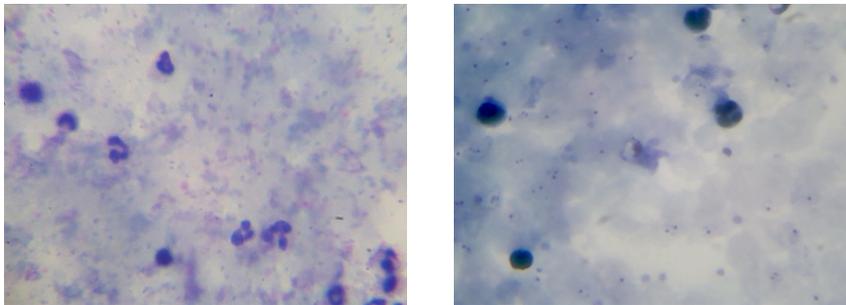
In order to automate the process of parasite detection, we first consider two different methods for capturing images of blood smears under a microscope: using a dedicated microscope camera, or using the camera of a mobile device such as a smartphone. For the latter method, we investigated the potential of 3D printing for producing low-cost adapters with which to mount a phone directly on the eyepiece of the microscope. The promise of this type of 3D printing approach is that customized adapters could in principle be made on demand for any combination of camera phone and microscope, as long as the geometry of the phone and the eyepiece diameter of the microscope eyepiece are known. Figure 8.1 (left) shows the design of our prototype adapter for attaching a ZTE Blade low cost Android smartphone to a Brunel SP150 microscope. Combining the imaging and computation on a single device, particularly a device already widely available even in malaria-endemic regions, would clearly make the system quite practical to deploy. Figure 8.1 (right) shows the printed adapter on one eyepiece, and a Motic MC1000 microscope camera on the other eyepiece.

Samples of the images taken from the camera phone and the dedicated microscope camera are shown in Figure 8.2. The image from the camera phone is clear, but has a wider field of view than the dedicated microscope camera. A single parasite is about 20 pixels across in the Motic image, but only around 8 pixels across in the image taken with the phone camera. We concluded that the phone imaging setup is a promising, low cost method for capturing blood film images for diagnosis, but that more work is needed on building extra magnification into such adapters in order to obtain a sufficient level of detail on plasmodium objects. The experiments described in the following parts of this chapter all use images from the Motic camera.

Sets of images were taken of thick blood smears with Field's stain from 133 individuals, using 1000x magnification with an oil immersion objective lens. After eliminating images which were out of focus, having motion artifacts from movement of the microscope stage during image capture, or were hyperparasitemic to the extent that it would be impractical to confidently label every parasite in the image, we were left with a set of 2703 images.



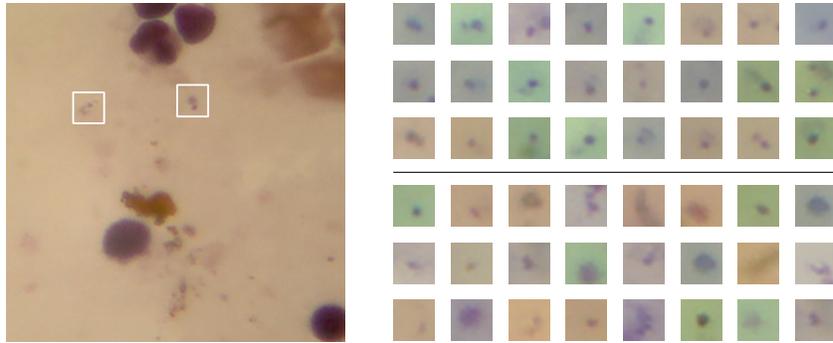
**FIGURE 8.1**  
3D-printable design for smartphone adapter (left). Printed phone adapter and Motic camera mounted on microscope (right).



**FIGURE 8.2**  
Image from camera-phone (left), and image from Motic camera (right).

#### 8.4.1 Annotation

In order to train and test the automated diagnosis system, it was necessary to annotate each of these images with the bounding boxes around each parasite. Bounding boxes were annotated on the captured images using labelling software developed for the PASCAL Visual Object Classes challenge [7]. A team of four experienced laboratory technicians used this software to indicate the position of every object they judged to be a parasite. Sample annotations are shown in Figure 8.3 (left). In this way the coordinates of 50,255 parasites were recorded within the set of captured images.

**FIGURE 8.3**

Bounding boxes around parasites annotated on a training image (left). Sample image patches close to the decision boundary, positive cases at the top and negative cases at the bottom (right).

---

## 8.5 Automated diagnosis

We split up each image into overlapping patches, and assigning each patch a label of 0 or 1, depending on whether the centre of a parasite bounding box is within that patch. Each  $1024 \times 768$  image was split into 475 overlapping patches, each of size  $50 \times 50$  pixels. Given this labelled set of image patches, we can pose the plasmodium detection task as a classification problem. To illustrate the nature of this problem, and its difficulties, we show examples of image patches in Figure 8.3 (right). The patches in the upper section of the Figure all have a positive label (i.e. they contain the centre of a parasite bounding box as specified by one of the expert annotators). The lower patches all have negative labels (i.e. they do not contain the centre of a bounding box), but contain artifacts, platelets or other shapes which might appear to a classifier to be close to the decision boundary.

The raw form of the pixel data in these image patches is not directly very useful for classification. We instead require a representation which is invariant with respect to rotation, translation and constant offsets in intensity. We may also require scale invariance if the images are not collected with a fixed magnification. Since the patches in the parasite recognition problem contain plasmodium against a background of normal cell matter, it also helps in this case to engineer features with some concept of an “object”. This is common in biomedical imaging applications, where we wish to segment images into objects such as organs, cells, and so on.

Feature engineering is therefore a significant step in the development of the automated diagnosis system. We have two aims in this: firstly to find a representation of the data which leads to good performance on the plasmodium

detection task, and secondly to have a sufficiently general representation of the shapes in blood smear images that other objects of interest—such as different hemoparasites, or white blood cells—could also be effectively identified in future with the same platform.

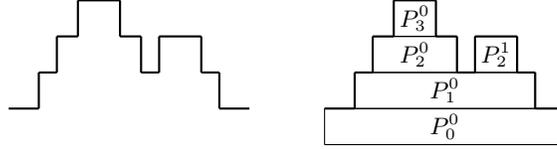
The plasmodium detection problem primarily concerns the shape of objects in the input patches. In general color information can also be useful (e.g. with Giemsa or Leishman stains), though it is not so informative using blood films treated with Field’s stain. Hence the features we use for this task are statistical representations of the shapes found in the image patches, and for all feature extraction we first convert the color patches to grayscale. We use two types of features: those derived from *connected components*, using concepts from mathematical morphology, and those derived from calculating moments of the patches thresholded at multiple levels. These two types of features are explained in Sections 8.5.1 and 8.5.2 respectively. We then conclude the automated diagnosis methodology by describing the classification process in Section 8.5.3.

### 8.5.1 Connected component features

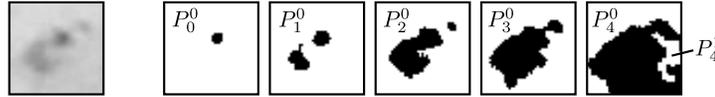
In this section we describe features based on regions, which are spatially connected sets of pixels that have some property in common such as similar grey level, and are used to define disjoint image segments. Regions, if properly defined, should correspond to objects. However, proper definition of regions is a difficult problem in image analysis. One approach to this problem, which has been used extensively in medical imaging, is a class of operators called connected filters [26, 10]. These are a family of morphological operators [9] that are based on the notion of connectivity and operate by interacting with connected components rather than individual pixels. Connectivity describes the way pixels are grouped to form connected components or flat zones in gray scale.

To calculate shape features for an image patch in this way, we first thresholded the image at each gray level. Connectivity openings [23] were used to calculate all the components in each thresholded image. These are known as peak components and denoted as  $P_h^k$  for gray level  $h$  and index  $k$ . This is illustrated in Figure 8.4 for a 1D example, and Figure 8.5 for an example with an image patch containing a parasite. The peak components were used to construct a max tree [25], which is a data structure designed for morphological image processing in order to efficiently compute features or attributes of the connected components. The max tree makes it possible to compute a large number of shape attributes for each of its nodes, and classification can then be based on these computed properties.

We computed several common morphological features for the connected components of every image patch, then used feature selection to narrow these down to five features most informative for the parasite detection task. These selected features are as follows:



**FIGURE 8.4**  
A 1-D signal  $f$  (left), and the corresponding peak components (right).



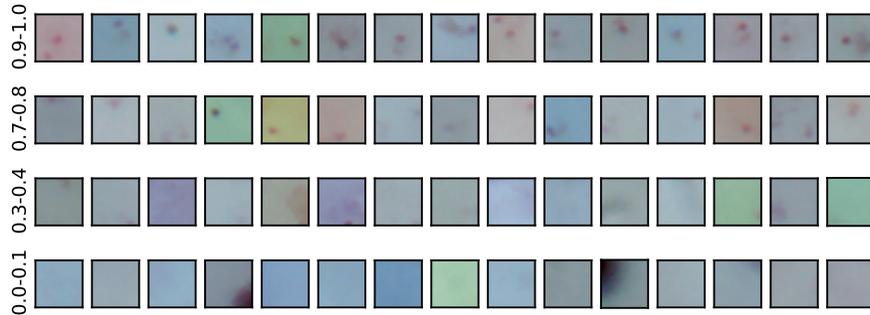
**FIGURE 8.5**  
Sample grayscale patch containing a parasite (left), and connected peak components—indicated as white regions—at different threshold levels (right).

- **Perimeter**
- **Moment of Inertia**
- **Elongation:**  $\text{Inertia}/\text{Area}^2$
- **Jaggedness:**  $(\text{Area} \times \text{Perimeter}^2) / (8\pi^2 \times \text{Inertia})$
- **Maximum  $\lambda$ :** Maximum child gray level - current gray level

Our aim was to compute standardized feature vectors for each patch in an image. Because the number of components on a single patch is variable, we summarized the shape information by traversing the max tree and calculating the percentile distribution of every attribute. For the five features above, we calculated the 25th, 50th, 75th percentiles, and the minimum and maximum values. Therefore we obtained a 25-dimensional feature vector for each image patch.

### 8.5.2 Moment features

An additional set of features was obtained by thresholding each patch at five different levels between the minimum and maximum pixel value. For each of the binary images returned by the threshold operation, we calculated many standard moment statistics [29] and used feature selection methods on training data to find those which were most discriminative with respect to the patch label. In this way, we selected the moment  $m_{00}$ , the central moments  $\mu_{11}$ ,  $\mu_{20}$ ,  $\mu_{02}$  and Hu moments  $h_0$ ,  $h_1$ ,  $h_2$ . These seven statistics, calculated for five different thresholded version of the patch, therefore provided an additional 35 features for each patch. These features were appended to those calculated as described in the previous section.

**FIGURE 8.6**

Sample patches in test images, arranged by the probability assigned by the classifier to each patch of its containing a parasite. The probability range of each row is indicated on the left, with the top row being patches confidently classified as positive cases, and the bottom row being confidently classified as negative cases.

### 8.5.3 Classification

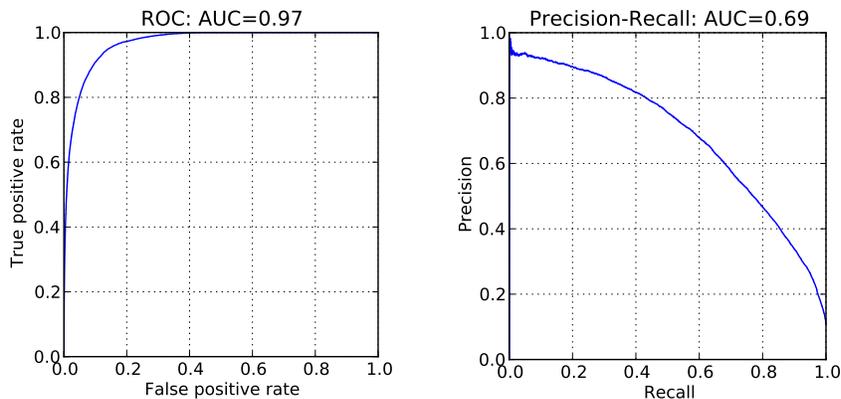
The Extremely Randomized Trees classifier [8] was used to learn a mapping between features and patch labels. This is a type of ensemble method, in which many decision trees are learned by selecting thresholds at random and retaining the trees which give good classification performance. Its advantages for this application, as well as good discriminative performance, are that it is fast and memory-efficient to evaluate at test time. This is useful for situations where classification is to be carried out on a mobile device with limited computational resources. An ensemble of 250 trees, with a maximum tree depth of 5 was used.

---

## 8.6 Evaluation

The classifier was trained on 75% of the labeled data (2027 images, containing 37550 patches annotated as containing parasites), and tested on the remaining 25% (676 images, containing 16312 patches annotated as containing parasites).

The receiver operating characteristic (ROC) curve is shown in Figure 8.7 (left), and the area under the curve (AUC) of 0.97 indicates that the classifier is effective at distinguishing positive and negative patches. Note that this is the performance at determining whether individual image patches contain parasites, not the performance at classifying all the image data from a single patient. If there is one or more positive patches within the set of images from

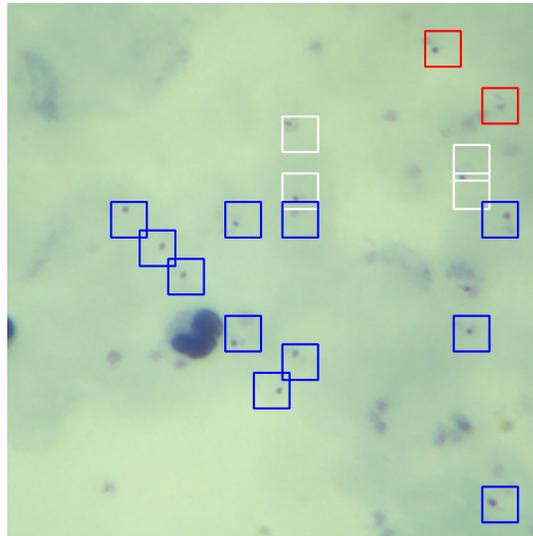


**FIGURE 8.7**  
Receiver Operating Characteristics and Precision-Recall curves for test data.

an individual sample, then that sample is considered infected. We cannot give per-patient sensitivity and specificity results here, as nearly all of the blood smear images in our experiments were from malaria-infected individuals.

The precision-recall curve, shown in Figure 8.7 (right) shows the different trade-offs possible between increasing sensitivity and decreasing the false alarm rate. Note that this compares favorably to the performance of related methods for thin blood smear analysis in Section 8.3.2, given that thick blood smears are an order of magnitude more sensitive than thin blood smears. Therefore, if we choose a detection threshold with gives us precision of 90%, the corresponding recall of around 20% is still higher than any method using thin blood smears would be able to attain after analyzing the same number of fields of view.

From Figure 8.7 (right) we can see that if the system were to be used for entirely automated diagnosis, in order to have a false alarm rate below one in ten, the recall would be around one fifth of what an experienced laboratory technician would be able to achieve; that is, the minimum detectable concentration of parasites in the blood would be around five times higher for the automated system than for the human expert. This is better than a purely symptomatic diagnosis, but clearly has shortcomings in terms of sensitivity. An alternative way to use such a system, therefore, would be as a decision-making aid to a technician. The aim in this context is to process the images from the microscope in order to focus the technician's attention on only the objects within those images that most resemble plasmodium. For this, a different threshold with greater sensitivity would be appropriate. For example, we might require recall of 90%, giving a corresponding precision of 37%. That is, we would expect the system to detect nine of every ten parasites appearing in the images, but 63% of detections would be false alarms. This would be

**FIGURE 8.8**

Sample detection output on part of a test image. Blue squares indicate patches which were correctly classified as containing parasites (where the classifier gave a probability above a threshold of 0.8, and the patch label was positive), red squares indicate false positives (classifier probability above threshold, but label negative) and white squares indicate false negatives (classifier probability below threshold, but label positive). By adjusting the classification threshold, different trade-offs are found between false positives and false negatives.

discriminative enough to considerably improve the throughput of a technician, who would only have to assess highlighted regions, rather than entire images.

The system was implemented in Python with scikit-learn<sup>1</sup> and opencv<sup>2</sup>. Feature extraction code was implemented in C for computational efficiency. Video of a netbook deployment operating in a clinical setting can be seen online<sup>3</sup>.

---

<sup>1</sup><http://scikit-learn.org>

<sup>2</sup><http://opencv.org>

<sup>3</sup><http://aidevmakerere.blogspot.com/2012/08/live-testing-of-computer-vision-malaria.html>

## 8.7 Discussion

In this chapter we have given a methodology for automated diagnosis of malaria from blood smear images, including image capture, feature extraction and classification. The accuracy of the system currently makes it practical as decision making aid for laboratory technicians, by triaging attention to the parts of images most indicative of plasmodium. For fully automated diagnosis, in order to have a false alarm rate below 10%, the sensitivity would be around 20% of what a trained microscopists would be able to achieve with the same images. This is still likely to be more sensitive than an analysis based on thin blood smears, which has been the focus of most previous work on vision-based automated malaria diagnosis, and certainly more sensitive than rapid diagnostic tests.

A platform for the automation of diagnosis from blood smear images provides several interesting and useful directions for future work. A location-aware mobile device could look up its location on a risk map, in order to set a prior for diagnosis (discussed further in [17]). Automated diagnosis also provides a significant opportunity for data collection; if test results are stored centrally, then spatial patterns of malaria incidence could be inferred. Furthermore, the feature extraction and classification framework we have described is sufficiently general that a whole suite of diagnostic tests—e.g. for tuberculosis, worm infestations, or hemoparasites other than plasmodium—could feasibly be implemented using the same framework.

---

## 8.8 Acknowledgements

This work was funded in part by a grant from Microsoft Research. Annotation of blood smear images was carried out by Alfred Andama, Vincent Wadda, Steven Ikodi and Patrick Byanyima. 3D printing was done by Sandeep Patel and Kenan Pollack of OmusonoLabs, Kampala. Mark Everingham provided us with image annotation software.



---

## *Bibliography*

- [1] K. Abba, J.J. Deeks, P. Olliaro, C-M. Naing, S.M. Jackson, Y. Takwoingi, S. Donegan, and P. Garner. Rapid diagnostic tests for diagnosing uncomplicated p. falciparum malaria in endemic countries. *Cochrane Database Syst Rev*, 7, 2011.
- [2] L. Allen, J. Hatfield, G. DeVetten, J. Ho, and M. Manyama. Reducing malaria misdiagnosis: the importance of correctly interpreting Paracheck Pf. *BMC infectious diseases*, 11(1):308, 2011.
- [3] B.B. Andrade, A. Reis-Filho, A.M. Barros, S.M. Souza-Neto, L.L. Nogueira, K.F. Fukutani, E.P. Camargo, L.M.A. Camargo, A. Barral, A. Duarte, and M. Barral-Netto. Towards a precise test for malaria diagnosis in the Brazilian Amazon: comparison among field microscopy, a rapid diagnostic test, nested PCR, and a computational expert system based on artificial neural networks. *Malaria Journal*, 9:117, 2010.
- [4] J. Bailey, J. Williams, B.J. Bain, J. Parker-Williams, and P.L. Chiodini. Guideline: the laboratory diagnosis of malaria. *British Journal of Haematology*, 163(5):573–580, 2013.
- [5] V. Batwala, P. Magnussen, K.S. Hansen, F. Nuwaha, et al. Cost-effectiveness of malaria microscopy and rapid diagnostic tests versus presumptive diagnosis: implications for malaria control in Uganda. *Malaria Journal*, 10(1):372, 2011.
- [6] D.N. Breslauer, R.N. Maamari, N.A. Switz, W.A. Lam, and D.A. Fletcher. Mobile phone based clinical microscopy for global health applications. *PLoS One*, 4(7):e6320, 2009.
- [7] M. Everingham, L. Van Gool, C.K.I. Williams, J. Winn, and A. Zisserman. The PASCAL Visual Object Classes (VOC) Challenge. *International Journal of Computer Vision*, 88(2):303–338, 2010.
- [8] P. Geurts, D. Ernst, and L. Wehenkel. Extremely randomized trees. *Machine Learning*, 63(1):3–42, 2006.
- [9] H.J.A.M. Heijmans. *Morphological Image Operators*. Academic Press, Boston, 1994.
- [10] H.J.A.M. Heijmans. Connected morphological operators for binary images. *Computer Vision and Image Understanding*, 73:99–120, 1999.

- [11] B. Houwen. Blood film preparation and staining procedures. *Laboratory Hematology*, 6(1):1–7, 2000.
- [12] D.S. Ishengoma, F. Francis, B.P. Mmbando, J.P.A. Lusingu, P. Magistrado, M. Alifrangis, T.G. Theander, I.C. Bygbjerg, and M.M. Lemnge. Accuracy of malaria rapid diagnostic tests in community studies and their impact on treatment of malaria in an area with declining malaria burden in north-eastern Tanzania. *Malaria Journal*, 10:176, 2011.
- [13] M. Kiggundu, S.L. Nsohya, M.R. Kamya, S. Filler, S. Nasr, G. Dorsey, and A. Yeka. Evaluation of a comprehensive refresher training program in malaria microscopy covering four districts of Uganda. *The American Journal of Tropical Medicine and Hygiene*, 84(5):820, 2011.
- [14] O.A. Koita, O.K. Doumbo, A. Ouattara, L.K. Tall, A. Konaré, M. Diakit , M. Diallo, I. Sagara, G.L. Masinde, S.N. Doumbo, et al. False-negative rapid diagnostic tests for malaria and deletion of the histidine-rich repeat region of the hrp2 gene. *The American Journal of Tropical Medicine and Hygiene*, 86(2):194–198, 2012.
- [15] V.V. Makkapati and R.M. Rao. Segmentation of malaria parasites in peripheral blood smear images. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009.
- [16] L.M. Milne, M.S. Kyi, P.L. Chiodini, and D.C. Warhurst. Accuracy of routine laboratory diagnosis of malaria in the United Kingdom. *Journal of clinical pathology*, 47(8):740–742, 1994.
- [17] M.G. Mubangizi, C. Ikae, A. Spiliopoulou, and J.A. Quinn. Coupling spatiotemporal disease modeling with diagnosis. In *Proceedings of the International Conference on Artificial Intelligence (AAAI)*, 2012.
- [18] C.K. Murray, R.A. Gasser, A.J. Magill, and R.S. Miller. Update on rapid diagnostic testing for malaria. *Clinical Microbiology Reviews*, 21:97–110, 2008.
- [19] World Health Organization. Malaria light microscopy: creating a culture of quality. In *Report of WHO SEARO/WPRO workshop on quality assurance for malaria microscopy, Geneva*, 2005.
- [20] D. Payne. Use and limitations of light microscopy for diagnosing malaria at the primary health care level. *Bulletin of the World Health Organization*, 66(5):621, 1988.
- [21] M.E. Rafael, T. Taylor, A. Magill, Y-W Lim, F. Girosi, and R. Allan. Reducing the burden of childhood malaria in Africa. *Nature*, 444:39–48, 2006.

- [22] A. Roca-Feltrer, I. Carneiro, A. Schellenberg, and J.R.M. Estimates of the burden of malaria morbidity in Africa in children under the age of 5 years. *Tropical Medicine & International Health*, 13(6):771–783, 2008.
- [23] C. Ronse. Set-theoretical algebraic approaches to connectivity in continuous or digital spaces. *J. Math. Imag. Vis.*, 8:41–58, 1998.
- [24] N.E. Ross, C.J. Pritchard, D.M. Rubin, and A.G. Duse. Automated image processing method for the diagnosis and classification of malaria on thin blood smear. *Med Biol Eng Comput*, 44:427–436, 2006.
- [25] P. Salembier, A. Oliveras, and L. Garrido. Anti-extensive connected operators for image and sequence processing. *IEEE Trans. Image Proc.*, 7:555–570, 1998.
- [26] P. Salembier and J. Serra. Flat zones filtering, connected operators, and filters by reconstruction. *IEEE Trans. Image Proc.*, 4:1153–1160, 1995.
- [27] E. M. Samba. The burden of malaria in Africa. *Afr. Health*, 19(2):17, 1997.
- [28] N. Tangpukdee, C. Duangdee, P. Wilairatana, and S. Krudsood. Malaria diagnosis: a brief review. *The Korean Journal of Parasitology*, 47(2):93–102, 2009.
- [29] C.-H. Teh and R.T. Chin. On image analysis by the methods of moments. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 10(4):496–513, 1988.
- [30] F.B. Tek, A.G. Dempster, and I. Kale. Malaria parasite detection in peripheral blood images. *British Machine Vision Conference*, pages 347–56, 2006.
- [31] F.B. Tek, A.G. Dempster, and I. Kale. Computer vision for microscopy diagnosis of malaria. *Malaria Journal*, 8:153, 2009.
- [32] F.B. Tek, A.G. Dempster, and I. Kale. Parasite detection and identification for automated thin blood film malaria diagnosis. *Computer Vision and Image Understanding*, 114(1):21–32, 2010.
- [33] D. Tseng, O. Mudanyali, C. Oztoprak, S.O. Isikman, I. Sencan, O. Yaglidere, and A. Ozcan. Lensfree microscopy on a cellphone. *Lab on a Chip*, 10(14):1787–1792, 2010.
- [34] M. Tumwebaze. Evaluation Of The Capacity To Appropriately Diagnose And Treat Malaria At Rural Health Centers In Kabarole District, Western Uganda. *Health Policy and Development*, 9(1):46–51, 2011.

- [35] H.A. Williams, L. Causer, E. Metta, A. Malila, T. O'Reilly, S. Abdulla, S.P. Kachur, and P.B. Bloland. Dispensary level pilot implementation of rapid diagnostic tests: an evaluation of RDT acceptance and usage by providers and patients—Tanzania, 2005. *Malaria Journal*, 7(1):239, 2008.
- [36] C. Wongsrichanalai, M.J. Barcus, S. Muth, A. Sutamihardja, and W.H. Wernsdorfer. A review of malaria diagnostic tools: microscopy and rapid diagnostic test (RDT). *The American Journal of Tropical Medicine and Hygiene*, 77(6 Suppl):119–127, 2007.