Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 90 (2016) 138 – 144

Procedia Computer Science

International Conference On Medical Imaging Understanding and Analysis 2016, MIUA 2016, 6-8 July 2016, Loughborough, UK

Automated detection of malaria parasites on thick blood smears via mobile devices

Luís Rosado[a,*], José M. Correia da Costa[b], Dirk Elias[a], Jaime S. Cardoso[c]

[a] *Fraunhofer Portugal AICOS, Rua Alfredo Allen 455/461, 4200-135 Porto, Portugal*
[b] *Instituto Nacional de Saúde Dr. Ricardo Jorge, Rua Alexandre Herculano 321, 4000-055 Porto, Portugal*
[c] *INESCTEC and University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal*

Abstract

An estimated 214 million cases of malaria were detected in 2015, which caused approximately 438 000 deaths. Around 90% of those cases occurred in Africa, where the lack of access to malaria diagnosis is largely due to shortage of expertise and equipment. Thus, the importance to develop new tools that facilitate the rapid and easy diagnosis of malaria for areas with limited access to healthcare services cannot be overstated. This paper presents an image processing and analysis methodology using supervised classification to assess the presence of *P.falciparum* trophozoites and white blood cells in Giemsa stained thick blood smears. The main differential factor is the usage of microscopic images exclusively acquired with low cost and accessible tools such as smartphones, using a dataset of 194 images manually annotated by an experienced parasilogist. Using a SVM classifier and a total of 314 image features extracted for each candidate, the automatic detection of trophozoites detection achieved a sensitivity of 80.5% and a specificity of 93.8%, while the white blood cells achieved 98.2% of sensitivity and 72.1% specificity.

Peer-review under responsibility of the Organizing Committee of MIUA 2016

Keywords: Image analysis; Malaria; Computer-aided diagnosis; Microscopy; Mobile Devices;

1. Introduction

Malaria is a leading cause of death and disease in many developing countries [1]. In 2015, there were an estimated 214 million cases of malaria, which caused approximately 438 000 deaths. Around 90% of malaria cases occurred in Africa, where the lack of access to malaria diagnosis is largely due to a shortage of expertise, the shortage of equipment being the secondary factor. In a recent report [1], the WHO considers that the current funding distribution of malaria control commodities (US$ 1.6 billion in 2014) is not addressing the fundamental weaknesses in health systems of developing countries, suggesting that innovative ways may be required to rapidly expand access to malaria interventions. It is worth underlining that the mobile phone is currently Africa's most important digital technology [2], and just as African telecommunications largely skipped over landline infrastructure and went straight to mobile phones, some experts say African medicine can skip over centralized labs [3]. Moreover, the combination of mobile devices

* Corresponding author. Tel.: +351 22 0430 336.
E-mail address: luis.rosado@fraunhofer.pt

Peer-review under responsibility of the Organizing Committee of MIUA 2016
doi:10.1016/j.procs.2016.07.024

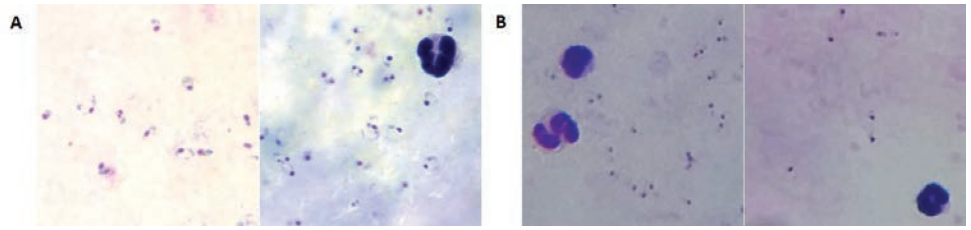

Fig. 1. Cropped microscopic sub-images of *P.falciparum* trophozoites and WBC on thick blood smear acquired with: (a) proper microscopic equipment[4]; (b) smartphone coupled to a Optical Magnification Prototype (see Section 3).

with image processing for malaria diagnosis can bring several advantages, like potentially reducing the dependence of manual microscopic examination, which is an exhaustive and time consuming activity, simultaneously requiring considerable expertise of the laboratory technician. Microscopy examination remains the gold standard for laboratory confirmation of malaria[4], which can be made through microscopic examination of thin and thick blood smears. While the thin smear consists in a single layer of red blood cells, the thick smear is 6-20 times thicker, allowing for a greater volume of blood to be examined. Thus, thick smears are firstly used to check the presence of malaria parasites (MP), while thin smears are subsequently analyzed for the identification of MP species.
This paper presents an image processing and analysis methodology using supervised classification to assess the presence of *P.falciparum* trophozoites and White Blood Cells (WBC) in Giemsa stained thick blood smears, using microscopic images acquired exclusively with smartphones.

2. Related Work

Image processing approaches have been proposed in order to identify parasites in malaria-infected thick blood smears[5,6,7,8,9]. *Kaewkamnerd et al.*[5] used an adaptive threshold computed from the V-Value histogram using only 20 images, thus requiring further validation. *Elter et al.*[6] have used a SVM classifier with a RBF kernel to identify objects containing chromatin, with reported sensitivity of 97%. Also, co-occurrence matrix and wavelet transform have been used by *Yunda et al.*[7], for detection of *P.vivax* in thick blood films, combined with the usage of Neural Networks. In a different approach, *Purnama et al.*[8] used Genetic Programming to detect different species and stages, with reported accuracy of 96% using 180 manually cropped sub-images. More recently, *Quinn et al.*[9] have extracted moment and connected component features for classification with Extremely Randomized Trees with stated results of AUC=0.97 for the ROC curve, being the only work found that uses smartphone-acquired images. Indeed, promising results for malaria diagnosis. Nevertheless, most of these methodologies are based on two criteria: i) images acquired under well controlled conditions; ii) the need of proper microscopic equipment. Both criteria are difficult to accomplish in endemic areas of Malaria, where this type of equipment is scarce or nonexistent in healthcare facilities. Alternatively, here, we present a different methodology approach for image processing of malaria-infected thick blood smears by using images exclusively acquired with low cost and accessible tools such as smartphone (see Fig. 1). The list of

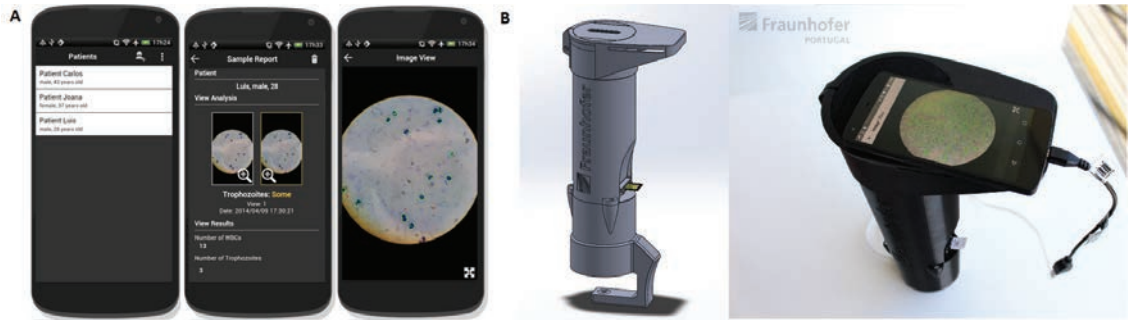

Fig. 2. Mobile-based Framework for Malaria Parasites Detection: (a) Smartphone Application; (b) Optical Magnification Prototype.

image features used in this work merges the outcomes of an extensive review of all the significant features proposed on the literature for MP and WBC detection [10], with new proposed features.

3. Mobile-based Framework for Malaria Parasites Detection: An Overview

The work reported in this paper represents only a component of a mobile-based framework for MP detection currently being developed, which is composed by 3 main components (see Fig. 2): **1) Optical Magnification Prototype**: a cheap alternative to the current microscopes, that can easily be adapted to a smartphone and used in the field. This gadget guarantees the required 1000x magnification, and the smartphone camera is used to capture images. Moreover, it uses a self-powered motorized automated stage system, in order to move the blood smear and allow the automatic capture of several snapshots of the sample; **2) Image Processing and Analysis**: consist of the automated detection of MP in blood smears through computer-aided methods. Four species of *Plasmodium* can infect and be transmitted by humans, passing through 3 different growth stages. Thus, there are 12 different species-stage combinations, for each we aim to develop image processing modules to identify and count the respective MP. **3) Smartphone Application**: envisioned to be used by technical personnel without specialized knowledge in malaria diagnosis. The user collects and prepares a blood sample of the patient, introducing it in a slot of 1). Using the companion mobile application, installed in the smartphone that is coupled to 1), the user can take pictures of the blood smear using the smartphone's camera, being subsequently analyzed by 2), so the correct procedures and medication can be administered.

4. Methodology

The proposed methodology can be divided into 3 main block: **Optical Circle Detection**; **WBC Detection**; and **Trophozoites Detection**. For the Optical Circle Detection, the image was firstly converted to grayscale and a median filter with large window factor was applied to eliminate the inner structures. The optical circle segmentation was performed using the Otsu's Method, a well-known histogram shape-based image thresholding routine. The remaining inner structures inside the optical circle were removed using a flood fill algorithm (see Fig. 4.B).
A dataset of microscopic images acquired from 6 different thick blood smears infected with *P.falciparum* was used for development and testing of the proposed approach. The images were acquired using the Optical Magnification Prototype (see Section 3), coupled to a smartphone. Two different smartphones were used, a HTC One S and a LG Nexus 5, with image resolutions ranging from 1456x2592 to 1944x2592 pixels. A total of 194 images were manually annotated by an experienced parasilogist from the Infectious Diseases Department of Instituto Nacional de Saúde Dr. Ricardo Jorge, with 352 identified trophozoites and 1935 white blood cells (see Fig. 3).

4.1. WBC Detection

There are several types of WBC, with different morphological characteristics, but every WBC has a nucleous surrounded by cytoplasm. When a thick smear is prepared, the cytoplasm membrane is destroyed, thus the WBC nucleus being only visible. The WBC detection is divided into 2 main tasks: **1) WBC Nuclei Candidates Segmentation**: segment all the elements on the image that are possible candidates of WBC nucleus, which stain blue to almost black with Giemsa stain. To detect the WBC candidates, a pre-processing technique called mean shift filtering was firstly

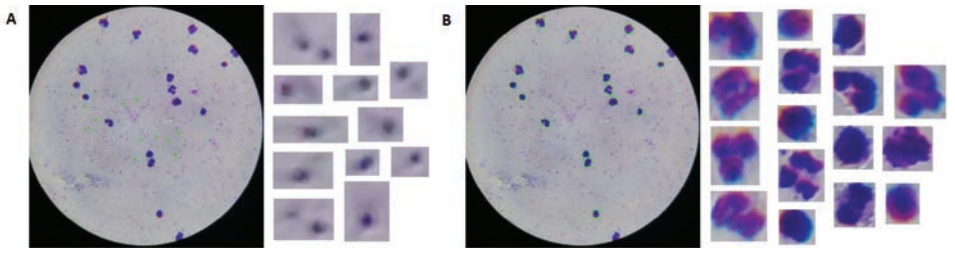

Fig. 3. Microscopic Image Dataset: (a) Trophozoites manual annotation; (b) White Blood Cells manual annotation.

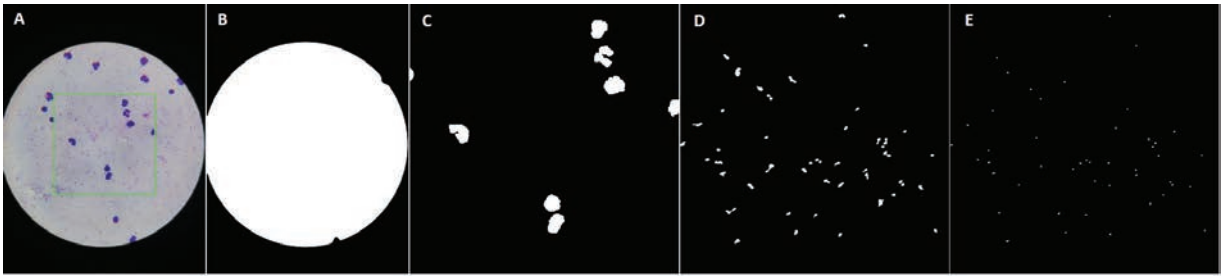

Fig. 4. Segmentation Results: (a) Original image with region of interest (ROI) at green; (b) Optical Circle segmentation mask; (c) WBC candidates segmentation mask of ROI; (d) Trophozoites candidates segmentation mask of ROI; (e) Chromatine dots candidates segmentation mask of ROI.

applied, using a spatial window radius of 5 and a color window radius of 15[11]. This edge-preserving smoothing method showed to facilitate the segmentation, as it preserves the edges of the WBC while the color inside the WBC nuclei becomes more homogeneous. An adaptive thresholding is then applied to segment the WBC candidates. Considering an original image I_L, the corresponding segmented I_S obtained by adaptive thresholding is given by Eq. 1. T_L is the mean intensity value of the square region centered on the pixel location (x,y) with a side value of R_S minus the constant C. In the proposed approach, it is used C=40 and R_S is given by Eq. 2, with a value for the division factor D_F=150. Furthermore, the Optical Diameter value represents the maximum diameter of the optical circle mask obtained previously. Since the microscopic images for malaria diagnosis will have a fixed magnification of 1000x, using the Optical Diameter as a metric reference turns this adaptive segmentation approach independent of the image resolution. Finally, the optical circle mask is subtracted and the final WBC candidates mask is obtained (see Fig. 4.C).

$$I_S(x,y) = \begin{cases} 0 & \text{if } I_L(x,y) > T_L(x,y) \\ 255 & \text{otherwise} \end{cases} \quad (1) \qquad R_S = \frac{\text{Optical Diameter}}{D_F} \quad (2)$$

It should also be taken into account that WBC nucleus can be composed by a single lobe or multiple detached lobes (see Fig. 3). To address the issue of detached lobes belonging to the same WBC, the WBC-candidates with size below the Optical Diameter are labelled as lobe-candidates. Through the expansion of the bounding box of every candidate, if the bounding box of the lobe-candidate intersects with the bounding box of either another lobe-candidate or a WBC-candidate, the bounding boxes are merged and the enclosing objects are labelled as a single WBC candidate. Otherwise, the lobe-candidates are discarded from classification consideration. It worth noting that lobe-candidates may intersect with several neighbouring bounding-boxes, so the Jaccard index was used as the tie-breaking criterion; **2) WBC Nuclei Candidates Feature Extraction and Classification:** For each WBC candidate, a total of 152 image features were extracted, grouped into 3 major groups: Geometry; Color; and Texture features (see Table 1). Moreover, for machine-learning training purposes, each candidate was labelled according to the manual annotation, i.e. labelled as WBC if the overlap coefficient (also known as Szymkiewicz-Simpson coefficient) between the region of interest of the candidate and the manual annotation is higher than 0.75. A two-class SVM classifier[12] with a RBF kernel was used to create a classification model, using a grid-search approach to obtain the best γ and C parameters.

4.2. Trophozoites Detection

Trophozoites are composed by a cytoplasm and one or two small chromatin dots. Using Giemsa stain, the cytoplasm usually stains blue and takes different shapes, from a well-defined, fine ring to forms that are irregular or bizarre, sometimes called amoeboid[13] (see Fig. 1). The Trophozoites Detection is divided into 3 main tasks: **1) Trophozoites Cytoplasm Candidates Segmentation:** The adaptive thresholding approach explained previously for the WBC segmentation was adapted for the segmentation of trophozoites' cytoplasm. Since the expect dimensions and stain contrast of trophozoite' cytoplasm is significantly lower when compared to WBC, the thresholding parameters were adapted through experimental testing (C=6, D_F=40). An area filtering process is then applied using Eq. 2, with D_F=25 and D_F=4 for minimum and maximum area, respectively; **2) Chromatine Dots Candidates Segmentation:** The chromatine dot is a part of the parasite nucleus, usually round in shape and stains red with Giemsa stain. It usually appears in the acquired images as a sharp small dot, so the algorithm consists on creating a local differences mask.

Table 1. Summary of the geometric, color and texture features extracted.

Group	Family	Channels	Features
Geometry		Binary	Area[a], Maximum Diameter[a], Minimum Diameter[a], Perimeter[a], Convex Hull Area[a], Solidity, Elongation Bounding Box Area[a], Extent, Equivalent Diameter[a], Circularity, Elliptical Symmetry, Radial Variance, Compactness Index, Principal Axis Ratio, Bounding Box Ratio, Irregularity Indexes, Eccentricity, Asymmetry Indexes, Asymmetry Ratios, Lengthening Index, Asymmetry Celebi
Color		C* and h (from L*C*h)	Energy[b], Mean[b], Standard Deviation[b], Entropy[b], Skewness[b], Kurtosis[b], L1 Norm[b], L2 Norm[b]
Texture	DFT	Grayscale	Mean, Standard Deviation, Minimum, Maximum
	GLRLM	Grayscale	Short run emphasis[c], Long run emphasis[c], Grey level non-uniformity[c], Run percentage[c], Low grey level runs emphasis[c], High grey level runs emphasis[c], Short run low grey level emphasis[c], Short run high grey level emphasis[c], Long run low grey level emphasis[c], Long run high grey level emphasis[c]
	GLCM	R, G, B (from RGB)	Energy[c], Entropy[c], Contrast[c], Dissimilarity[c], Homogeneity[c], Correlation[c], Maximum probability[c]
	Laplacian	Grayscale	Mean, Standard Deviation, Minimum, Maximum

[a] Feature divided by the Optical Diameter (used as metric reference), in order turn it independent of image size.; [b] Feature computed independently for each channel, as well as for the grayscale masks that result by folding and subtracting the region of interest of each channel according the major and minor axis of inertia.; [c] Feature computed for the following directions: 0^o, 45^o, 90^o and 135^o.

The original image is blurred using a median blur filter with a window size calculated using Eq. 2 with D_F=80. The small structures will disappear with the blurring process, among them the chromatine dots. The original image is then subtracted to the blurred image, and a 3-channel local differences mask is obtained. To binarize this mask, the maximum value in each pixel position along the 3-channels is selected, and the adaptive thresholding explained previously is applied with C=10 and D_F=80. To obtain the final chromatine dots mask, an area filtering process is then applied using Eq. 2, with D_F=125 and D_F=25 for minimum and maximum area, respectively. This mask is then used to filter the trophozoite' cytoplasm candidates mask, being only considered for further processing the cytoplasm candidates with at least one chromatine dot inside (see Fig. 4.D and 4.E); **3) Trophozoites Candidates Feature Extraction and Classification:** Each trophozoite candidate will be composed by a trophozoite' cytoplasm candidate and at least one chromatine dot candidate. A total of 314 features are extracted for each trophozoite candidate: the 152 image features referred on Section 4.1, extracted independently for the correspondent trophozoite' cytoplasm and for the chromatine dots candidates; 10 ratios between specific cytoplasm and chromatine features (area, maximum and minimum diameter, perimeter, convex hull area, bounding box area, relative difference of C* and h channels mean values, difference of the coefficient of variation of C* and h channels). Moreover, the same approach referred on Section 4.1 was applied for candidates labeling and machine learning classification.

5. Results

The classification results are presented in terms of three metrics: 1) Sensitivity, i.e. the percentage of candidates correctly classified as WBC/MP; 2) Specificity, i.e. the percentage of candidates correctly classified as Non-WBC/Non-MP; and 3) Accuracy, i.e. the percentage of candidates correctly classified overall. In Table 2 the results for WBC and *P.falciparum* trophozoites candidates after segmentation and filtering are depicted, while in Table 3 are detailed the results after machine-learning classification using candidate region-wise 10-fold cross-validation.

Table 2. Results for WBC and *P.falciparum* trophozoites candidates after segmentation and filtering.

	Manual Annotations	True Positives	False Positives	False Negatives
WBC	1935	1850	265	85
Trophozoites	352	277	1566	75

Table 3. Results for WBC and *P.falciparum* trophozoites detection after machine-learning classification.

	SVM Params	True Positives	True Negatives	False Positives	False Negatives	Sensitivity	Specificity	Accuracy
WBC	γ=0.2 C=1.0	1817	191	74	33	98.2%	72.1%	94.9%
Trophozoites	γ=0.03 C=0.025	223	1469	97	54	80.5%	93.8%	91.8%

The suggested methodology was implemented in C++ using the OpenCV library. In terms of performance, the registered average computational time was 4,59 seconds running on an Intel® Core™ i7-4790 CPU with 3.60GHz (with OS Ubuntu 14.04 LTS). The implemented C++ code was also deployed on a Samsung Galaxy 5 (CPU Quad-core 2.5 GHz Krait 400 running on Android OS 5.0), with an average time of 44 seconds. Moreover, the memory allocation of the proposed methodology was also analyzed using the Valgrind profiling tool, with maximum memory peaks detected always below 100MB for the images used in this study.

6. Conclusions and Future Work

In this work, a methodology to assess the presence of *P.falciparum* trophozoites and WBC in Giemsa stained thick blood smears is presented. The great majority of the proposed methodologies to date are based on images acquired under well controlled conditions and with proper microscopic equipment, so the main differential factor of this work is the usage of microscopic images exclusively acquired with smartphones coupled to a low cost optical magnification device, as well as the consequent customization of the proposed methodology for images with such characteristics. Given the lack of freely available image datasets, a mobile acquired image dataset manually annotated by a specialist was specifically created and used in this study. Moreover, a wide variety of image features were used to characterize the candidates in terms of geometry, texture and color, merging significant features referred on previous works, with some never used before for this purpose. In terms of results, the automatic detection of WBC in thick blood smears achieved 98.2% of sensitivity and 72.1% specificity, while the *P.falciparum* trophozoites detection achieved a sensitivity of 80.5% and a specificity of 93.8%.
It should be noted that this work represents only a component of a mobile-based framework for malaria parasites detection currently being developed. As future work in terms of image processing, we aim to develop methodologies for the analysis of thin blood smears images acquired exclusively with smartphones, in order to identify and count all possible species-stages combinations of MP that can potentially infect humans. The main final goal is to develop a system that can provide an effective pre-diagnosis of malaria to be used in medically underserved areas.

Acknowledgements

We would like to acknowledge the financial support from North Portugal Regional Operational Programme (NORTE 2020), Portugal 2020 and the European Regional Development Fund (ERDF) from European Union through the project 'Deus ex Machina: Symbiotic Technology for Societal Efficiency Gains', NORTE-01-0145-FEDER-000026.

References

1. World Health Organization | World Malaria Report. 2015. URL: `www.who.int/malaria/publications/world-malaria-report-2015`.
2. Blycroft Publishing | Africa & Middle East Mobile Factbook 2Q 2014. 2014. URL: `www.africantelecomsnews.com`.
3. Dolgin, E.. Portable pathology for Africa. *IEEE Spectrum* 2015;**52**(1):37–39.
4. Centers for Disease Control and Prevention. 2016. URL: `www.cdc.gov`.
5. Kaewkamnerd, S., Intarapanich, A., Pannarat, M., Chaotheing, S., Uthaipibull, C., Tongsima, S.. Detection and classification device for malaria parasites in thick-blood films. In: *2011 IEEE 6th International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS)*; vol. 1. 2011, p. 435–438.
6. Elter, M., Hasslmeyer, E., Zerfass, T.. Detection of malaria parasites in thick blood films. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 2011;**2011**:5140–5144.
7. Yunda, L., Alarcn, A., Milln, J.. Metodo automatizado de analisis de imagenes para deteccion del parasito de la malaria p-vivax en imagenes de gota gruesa. *Sistemas & Telematica* 2012;**10**(20):9–25.
8. Purnama, I., Rahmanti, F., Purnomo, M.. Malaria parasite identification on thick blood film using genetic programming. In: *3rd International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering*. 2013, p. 194–198.

9. Quinn, J., Andama, A., Munabi, I., Kiwanuka, F.. Automated Blood Smear Analysis for Mobile Malaria Diagnosis. In: *Mobile Point-of-Care Monitors and Diagnostic Device Design*; Devices, Circuits, and Systems. 2014, p. 115–132.
10. Rosado, L., Correia da Costa, J.M., Elias, D., S Cardoso, J.. A review of automatic malaria parasites detection and segmentation in microscopic images. *Anti-Infective Agents* 2016;**14**(1):11–22.
11. Wang, L., Liu, G., Dai, Q.. Optimization of Segmentation Algorithms Through Mean-Shift Filtering Preprocessing. *IEEE Geoscience and Remote Sensing Letters* 2014;**11**(3):622–626.
12. Chang, C.C., Lin, C.J.. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2011; **2**:27:1–27:27.
13. World Health Organization | Basic malaria microscopy: Part I. Learner's guide. Second edition. 2010. URL: `www.who.int/malaria/publications/atoz/9241547820`.