Sequential classification system for recognition of malaria infection using peripheral blood cell images

Angel Molina ,[1] Santiago Alférez,[2] Laura Boldú ,[1] Andrea Acevedo,[1,3] José Rodellar,[3] Anna Merino [1,4]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/jclinpath-2019-206419).

[1]Biochemistry and Molecular Genetics, Biomedical Diagnostic Center, Hospital Clinic de Barcelona, Barcelona, Spain
[2]School of Engineering, Science and Technology, Universidad del Rosario Facultad de Ciencias Naturales y Matemáticas, Bogota, Cundinamarca, Colombia
[3]Matemáticas CoDAlab, Universitat Politecnica de Catalunya, Barcelona, Catalunya, Spain
[4]Hospital Clinic de Barcelona, Barcelona, Catalunya, Spain

Correspondence to
Dr Angel Molina, Biochemistry and Molecular Genetics, Biomedical Diagnostic Center, Hospital Clinic de Barcelona, Barcelona 08036, Spain; molinaborras@gmail.com

Received 28 December 2019
Revised 21 February 2020
Accepted 25 February 2020
Published Online First 16 March 2020

To cite: Molina A, Alférez S, Boldú L, *et al*. *J Clin Pathol* 2020;**73**:665–670.

ABSTRACT

Aims Morphological recognition of red blood cells infected with malaria parasites is an important task in the laboratory practice. Nowadays, there is a lack of specific automated systems able to differentiate malaria with respect to other red blood cell inclusions. This study aims to develop a machine learning approach able to discriminate parasitised erythrocytes not only from normal, but also from other erythrocyte inclusions, such as Howell-Jolly and Pappenheimer bodies, basophilic stippling as well as platelets overlying red blood cells.

Methods A total of 15 660 erythrocyte images from 87 smears were segmented using histogram thresholding and watershed techniques, which allowed the extraction of 2852 colour and texture features. Dataset was split into a training and assessment sets. Training set was used to develop the whole system, in which several classification approaches were compared with obtain the most accurate recognition. Afterwards, the recognition system was evaluated with the assessment set, performing two steps: (1) classifying each individual cell image to assess the system's recognition ability and (2) analysing whole smears to obtain a malaria infection diagnosis.

Results The selection of the best classification approach resulted in a final sequential system with an accuracy of 97.7% for the six groups of red blood cell inclusions. The ability of the system to detect patients infected with malaria showed a sensitivity and specificity of 100% and 90%, respectively.

Conclusions The proposed method achieves a high diagnostic performance in the recognition of red blood cell infected with malaria, along with other frequent erythrocyte inclusions.

INTRODUCTION

Malaria is among the most important parasitic diseases in the world caused by different species of *Plasmodium* parasites. WHO reported 219 million cases of malaria in 2017, and over 435 000 deaths.[1] An early diagnosis is essential for effective disease management.[2] Thin peripheral blood smear (PBS) microscopic examination is a low-expensive and easily accessible tool for malaria diagnosis being the gold standard for species identification,[3] for parasite quantification and it is essential to detect cell abnormalities.[4 5] However, PBS examination is time consuming, needs expert professionals and is prone to interobserver variability.[6] Moreover, external quality programmes show that the presence of malaria in PBS samples is overlooked by some laboratories.[7] These drawbacks have contributed to develop image processing techniques that help to preclassify blood cells[8 9] contributing to more standardised morphological analysis.[10–12]

Machine learning methods have been proposed for malaria recognition.[13] Segmentation is the first step in machine learning approaches. Usually, it is done from the green component of RGB images, using histogram thresholding techniques[14–16] and applying watershed algorithm.[17–19] In other works, segmentation is based on granulometry analysis[20] or in the use of thresholding from the HSV colour space.[21] Segmentation is followed by feature extraction, which uses colour, textural and geometric morphological information.[22] Features are used to train classifiers, which in most cases are based on support vector machine (SVM),[19 23–25] k-nearest neighbours (KNN)[16 17 21 26] or decision trees algorithms.[27] Latest trends in classification of red blood cells (RBC) infected with malaria propose the use of deep learning approaches.[14 17 18 25 28] Literature reveals that the automatic recognition of malaria has been addressed: (1) to discern between infected versus non-infected RBC, (2) to differentiate between *Plasmodium* species or (3) to discriminate among different parasite stages. It is striking that none of the studies have considered other RBC inclusions such as Howell-Jolly (HJ) and Pappenheimer (PP) bodies, basophilic stippling (BS) or platelets (PLT) overlying RBC, which can confuse machine learning tools.

Evaluations of the Advanced RBC Application, a CellaVision software that performs RBC classification and recognises inclusions, reported sensitivity and specificity values of 23.5% and 81.1%, respectively.[29 30] Furthermore, malaria parasites are commonly misinterpreted as HJ bodies, PP bodies or BS. ARBC algorithms need further improvements, and require postclassification by clinical pathologist.

The objective of this work is to design a new system for the automatic recognition of malaria using digital image analysis and machine learning tools, which must be highly sensitive and specific for clinical application. This methodology aims to discriminate parasitised RBC not only from normal RBC, but also from other RBC inclusions, such as HJ bodies, PP bodies, BS as well as PLT overlying RBC.

MATERIALS AND METHODS

The proposed system is illustrated in figure 1. Inputs are digital RBC images acquired from PBS. After preprocessing, RBC are segmented and

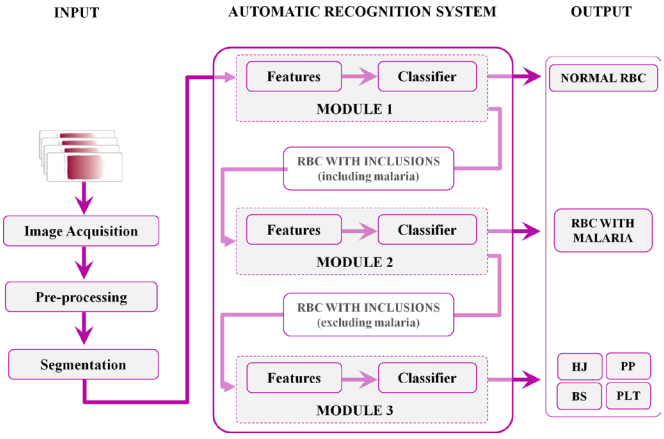

Figure 1 Illustration of the proposed system for the automatic classification of malaria and inclusions in red blood cells (RBC). The left track starts with the acquisition of images and ends up with the segmentation of the regions of interest, which are the inputs to the recognition system.The automatic recognition system includes three sequential modules. The first module recognises normal RBC with respect to those containing inclusions (including malaria). The second identifies whether the inclusions correspond to malaria parasites or not. In a third step, non-malaria inclusions are recognised individually as Howell-Jolly bodies (HJ), pappenheimer bodies (PP), basophilic stippling (BS) or platelets (PLT) overlying RBC.

represented by individual masks. Images and masks are the inputs to the automatic recognition system, which has a hierarchical architecture with three modules working sequentially: (1) the first recognises normal RBC with respect to RBC containing inclusions (including malaria); (2) the second identifies whether the inclusion corresponds to malaria or not and (3) the third discriminates among HJ, PP, BS or PLT. The output of the overall system is the automatic classification of the RBC images in one of the six classes under study.

The core of each module is a set of quantitative features along with a machine-learning classification algorithm. The study is designed in three steps: (1) datasets and image processing; (2) development, where classifiers are trained with selected features and (3) assessment of the overall system performance. Algorithms were developed in the programming environment Python 3.6, with the exception of RBC labelling that was performed in the scientific software MATLAB R2017b.

Datasets and image processing

Image acquisition and labelling

Blood samples, collected in Ethylenediaminetetraacetic acid (EDTA) anticoagulant, from 87 patients were stained with May Grünwald-Giemsa using the slide maker-stainer SP1000i (Sysmex, Kobe, Japan).[31 32] Field images were acquired using the Olympus BX43 microscope and Olympus DP73 camera. Images were taken at 1000 magnifications, 12 bits of colour depth and 2400×1800 pixels. Images were split into two sets of different smears: a training set, to develop the whole system and an assessment set to evaluate the system. This split was made ensuring that both datasets contained images from different patients.

Image labelling was performed by clinical pathologists, using Image Labeller app from MATLAB. It allows to select erythrocytes from field images, providing a marked area that associates the ground truth (true class) and location of every RBC. We considered the following types of images: (1) normal RBC; (2) RBC showing malaria parasites and (3) RBC containing other inclusions (HJ, PP, BS or PLT overlying them). Examples are shown in figure 2.

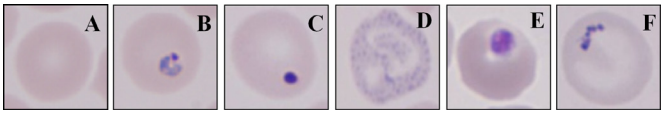

Figure 2 Example of different types of red blood cells (RBC) considered in the study. (A) Normal RBC. (B) RBC infected with malaria parasites. (C) RBC containing Howell-Jolly bodies. (D) RBC with basophilic stippling. (E) RBC with platelets allocated on its surphace. (F) RBC containing Pappenheimer bodies.

Image preprocessing and segmentation

Preprocessing aims to enhance image properties to facilitate further operations over the regions of interest (ROIs). The green component was selected from the original RGB image. Contrast between RBC and background was improved using contrast local adaptive histogram equalisation. Gaussian and median filter were used to smooth the image and decrease background noise.

Segmentation was the next step, which gives a separation of the ROIs. Since nucleus is absent in RBC, segmentation is simpler compared with leukocytes.[33] Image was binarised by thresholding using Otsu's method. Post-processing techniques were applied, such as removing noise and PLT, filling holes to deal with the RBC central pallor and clearing borders to select the complete RBC. Watershed method was used to separate adjacent cells, allowing to identify each RBC as an individual element.

Using coordinates stored in the previous labelling step, it was possible to crop individual RBC images and its mask from the original and segmented images.

Feature extraction

Feature extraction from the segmented images of the training set was done using PyRadiomics 2.2.0,[34] an open source platform which calculates different types of features: shape, colour (first order statistics) and texture (grey-level co-ocurrence matrix, grey-level run-length matrix, grey-level size zone matrix, neighbouring grey tone difference matrix and grey-level dependence matrix features). Features give quantitative information to characterise each ROI, facilitating the subsequent learning and classification steps. Since inclusions are found in RBC with different shapes, extraction was based on the calculation of colour and texture features. They were extracted from each component of different colour spaces included in the python library scikit-image: RGB (red,green,blue), HSV (hue, saturation, value), XYZ (tristimulus values), LAB (lightness from black to white, from green to red, and from blue to yellow), HED (haematoxylin, eosin, dab), YIQ (luminance, in-phase quadrature), LCH (luminance, chroma, hue), YUV (luminance, blue chroma projection, red chroma projection), YCbCr (luminance, blue-difference chroma, red-difference chroma)and CMYK (cyan, magenta, yellow, black), resulting in 2852 features calculated from each RBC.

System development

For each one of the three modules we followed the same procedure. The purpose was to train separately each classifier (see figure 1) to be highly specific for its particular recognition task. Essentially, a classifier uses a set of selected features to predict the class to which the image belongs. Several well-known machine learning algorithms were evaluated: SVM, KNN, linear

discriminant analysis (LDA), random forest and Gaussian Naive Bayes. For each classifier, all images of the training set were used to perform successive tests of fivefold cross-validation for different increasing numbers of features. Features obtained were ordered by relevance according to the conditional mutual information maximisation algorithm.[35] Selection of the best classification algorithm and the appropriate number of features was made according to highest overall accuracy, which is defined as the percentage of images well classified over the total number of images.

System assessment

The assessment set was divided into two sets (test and diagnostic) to perform an evaluation of the final trained recognition system through two approaches:

1. The test set was used to evaluate the performance to classify individual images. This set included images from 18 smears that were not used to develop the model.
2. The diagnostic set was used to assess the ability of the system to identify patients infected with malaria from the analysis of a whole PBS. A number of 18 new smears of different patients (not previously used) were included, eight of them corresponding to malaria patients showing different parasitaemia percentages. The remaining 10 were obtained from patients free of malaria, in which other RBC inclusions were observed.

RESULTS

Image processing and segmentation

Preprocessing and segmentation stages allowed to obtain images of individual RBC from microscopic field images. As an example, the different stages are illustrated in figure 3. The methodology used in this work showed a segmentation performance of 97.4%, considering a correct segmentation when the resulting mask preserved the original shape of its RBC. As a result, a total of 15 660 RBC images were obtained: 4633 RBC images for the training set and 11 027 RBC images for the assessment set (test and diagnostic sets). The total number of images for each RBC type is shown in table 1.

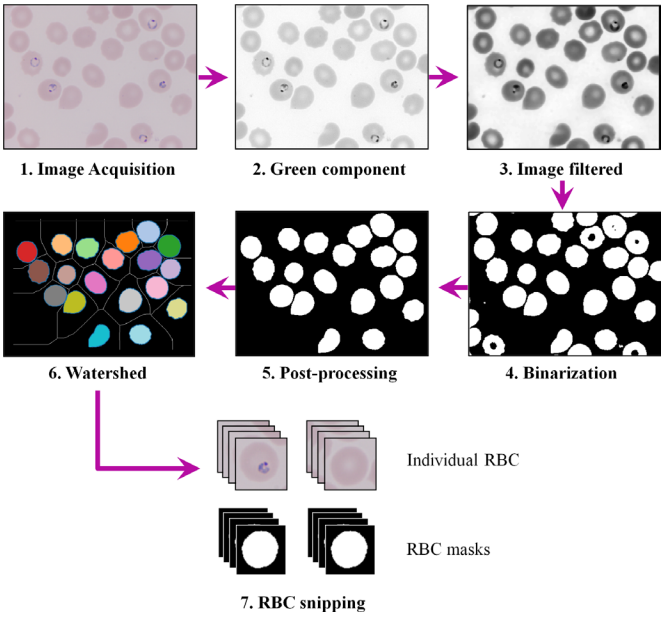

Figure 3 Illustration of the processing methodology used for the automatic classification of red blood cells (RBC), starting from the acquisiton of images and ending up with the obtention of individual RBC images and their binary masks.

Table 1 Total number of red blood cell (RBC) images and smears used in this work

		Image dataset			
			Assessment		
		Training	Test	Diagnostic	Total
No of smears		51	18	18	87
No of images	nRBC	1276	733	8443	10 452
	MAL	933	187	71	1191
	HJ	578	539	76	1193
	PP	698	256	72	1026
	BS	701	359	65	1125
	PLT	447	162	64	673
Total		4633	2236	8791	15 660

Images are grouped by class for each dataset (training set, test set, and diagnostic set).
BS, basophilic stippling; HJ, Howell-Jolly bodies; MAL, malaria parasites; nRBC, RBCs without inclusions; PLT, platelets; PP, Pappenheimer bodies.

System training

The final system combines three individual modules that work sequentially. Considering the discrimination among normal RBC and RBC containing any type of inclusion, we observed high accuracies with all methods along with low number of features. SVM, together with the seven most relevant features, exhibited the best accuracy (99.12%) using the training set. Regarding the second module, which recognises whether the RBC inclusions correspond to malaria parasites or not, LDA and SVM were the methods with highest accuracies. The combination of LDA with the 610 most relevant features showed accuracy values of 96.52%. LDA and SVM were the methods showing the highest accuracies for the third module to discriminate among HJ or PP bodies, BS and PLT overlying them. The highest accuracy (96.91%) was obtained for LDA along with the 700 most relevant features. Online supplementary figure 1 plots the accuracies with the five machine learning methods mentioned above as a function of different increasing numbers of features.

System assessment using the test set

Classification performance using the test set is summarised by the confusion matrix shown in table 2, which gives the sight of

Table 2 Confusion matrix of the classification results (in percentage and absolute values in brackets) using red blood cell (RBC) images of the test set

		Predicted class					
		nRBC	MAL	HJ	PP	BS	PLT
True class	nRBC	98.9 (725)	0	0	0	0.3 (2)	0.8 (6)
	MAL	0	100 (187)	0	0	0	0
	HJ	0	0.4 (2)	97.4 (525)	2.2 (12)	0	0
	PP	0	0.4 (1)	2 (5)	97.7 (250)	0	0
	BS	0.3 (1)	5.3 (19)	0	0	94.4 (339)	0
	PLT	0.6 (1)	1.2 (2)	0	0	0	98.1 (159)

Rows indicate the true class and columns represent the predicted class provided by the system. Diagonal values are the true positive rates.
BS, basophilic stippling; HJ, Howell-Jolly bodies; MAL, malaria parasites; nRBC, RBCs without inclusions; PLT, platelets; PP, Pappenheimer bodies.

Table 3 System evaluation of the diagnostic assessment set using individual smears from eight malaria-infected patients and 10 non-infected patients

		Predicted class		
	A	nRBC	MAL	INC
True class	nRBC	8439	0	4
	MAL	0	71	0
	INC	0	1	276

		Predicted class	
	B	Infected	Non-infected
True class	Infected	8	0
	Non-infected	1	9

A: confusion matrix of the classification results (in absolute number of images) for the diagnostic assessment of the recognition system. Rows indicate the true class and columns represent the predicted class provided by the system. Diagonal values are the true positive rates. B: confusion matrix for true and predicted diagnosis of individual smears tested. Smears are diagnosed as infected if the model recognises at least one single RBC image containing parasites.
INC, inclusions; MAL, malaria parasites; nRBC, normal red blood cells.

predicted vs true classes, in percentage and in absolute number of images. The main diagonal shows the sensitivity or true positive rate for each cell class (percentage of images predicted as positive over the total number of images that truly belong to the class): 98.9% for normal RBC, 100% for RBC infected with malaria parasites, 97.4% for HJ bodies, 97.7% for PP bodies, 94.4% for BS and 98.1% for PLT overlying RBC. The mean of these values (97.7%) is the overall accuracy, which is the percentage of single images that have been correctly recognised.

It is remarkable the 100% sensitivity achieved for the recognition of RBC infected with malaria versus all the non-infected RBC types. For this case, we may calculate the specificity as the ratio of images predicted as non-malaria (the sum of all columns in table 2 excluding malaria: 2025) over the number of images that truly belong to any of the non-malaria classes (the sum of all the rows in table 2 except for malaria: 2049). This gave a remarkable specificity of 98.8%.

System assessment for malaria infection using the diagnostic set

This evaluation was focused on the automatic recognition of malaria diagnosis using 18 smears corresponding to new patients whose images were not previously used. Table 3 presents the results for the 18 smears belonging to 8 patients with malaria, with different parasitaemia percentages (from 0.10% to 14.16%); and 10 non-malaria infected patients, but containing inclusions in some of their RBC. For each smear, the system classifies all the images in three classes: normal RBC, malaria or RBC containing other inclusions. The confusion matrices are shown in table 3A. Parasites were detected in all eight smears (71 RBC images with parasites), including those with low parasitaemia percentages. A patient was considered infected (predicted diagnosis) when at least one single RBC containing parasites was detected in the corresponding smear. Table 3B shows that sensitivity of the classification by smear was 100%.

We only found in one smear a single RBC image (with presence of an HJ inclusion) that was misidentified as malaria. According to the established diagnostic criterion, this smear was the only one (over 10) misdiagnosed as infected (90% of specificity). But considering individual RBC images, only one non-infected RBC image over a total of 8720 was misidentified as infected by malaria. Classification rates considering individual RBC images of the diagnostic set resulted into a sensitivity and a specificity of 100% and 99.98%, respectively. Online supplementary tables 1 and 2 show confusion matrices and diagnostic prediction for each specific smear.

DISCUSSION

Previous publications have addressed the malaria recognition problem.[13] Some of these works described machine learning approaches to recognise among different species of malaria or among different parasite stages.[22] However, these systems have only been designed to address the recognition between infected and non-infected RBC. In real practice, there are other inclusions that should be considered to obtain a sensitive and specific support tool. This remark is important because the recognition of infected samples is a critical result of urgent notification, since it can condition the patient's immediate treatment and prognosis.[36]

In this study, we developed a recognition system which is able to differentiate not only RBC infected with malaria, but also RBC containing other inclusions. To the authors' knowledge, it is the first time in the literature that this approach has been considered.

With respect to the segmentation of RBC and parasites, other works have used the histogram thresholding, euclidean distance transform,[37] edge detection algorithms[38] or fuzzy C means.[39] Those works that assess the segmentation performance show accuracies over 95%. We have obtained an overall segmentation accuracy of 97.4%. Our approach is based on Otsu's thresholding, which shows a good balance between implementation simplicity and performance.[13]

We evaluated different machine learning approaches for the design of three sequential classification modules, which showed high accuracy rates. Results showed that only a small number of features was required for the first classifier to discriminate between normal and RBC containing inclusions (including malaria). However, a higher number of features was necessary when addressing the recognition between malaria parasites and other inclusions (module two), or among different types of inclusions (module three). This is understandable because of the morphological similarities between the RBC images showing any type of inclusion.

SVM[19 23–25] and KNN[16 17 21 26] are among the most frequently used approaches to address the recognition of malaria parasites. In our work, SVM was also the best classifier for the first module. For modules two and three, LDA showed the best performance, while KNN showed lower accuracies. The methodological differences in the segmentation approach with respect to other publications, as well as the different techniques used for the extraction and analysis of features, may give the advantage of some models over the others. LDA has also been reported by our group to have a high accuracy in a machine learning system for the automatic recognition of different types of acute leukaemia.[40] Previous papers have reported accuracies ranging from 86% to 99%. However, they have not considered other RBC inclusions different from malaria. In our work, an overall accuracy of 97.7% was achieved in the assessment of the system by individual images, but addressing the recognition of malaria along with several other inclusions. This more complete differentiation is important in daily practice to ensure higher specificity in the automatic recognition of malaria.

It is important to consider individual patients when organising datasets for training and assessing the classifier. Using different images from the same smear in both training and assessment could result in an accuracy overestimation. To avoid this issue, we split the initial dataset into two sets of different smears. We only found two previous publications that followed this procedure when performing training/assessment splitting.[14 24] The classification of the test set images showed a very high performance: none of the normal RBC was classified as containing malaria and none of the infected RBC was classified as normal (sensitivity 100%). This ensures that patients with malaria do not go unnoticed. Besides, it is important to remark that the specificity of the proposed system is 98.8% in spite of the similarity of most of the RBC inclusions under study with respect to malaria parasites, which can easily confuse the classifier.

We also evaluated the ability of the system for the diagnosis of malaria infection on a different set of smears from new patients. The results showed that all the smears from the eight new patients containing RBC parasites were identified as infected. Although a single smear showing other RBC inclusions (due to a HJ body) was identified as infected, in practice, there was only one RBC image misclassified out of 8720 images well recognised as non-infected. These results showed that the approach proposed in this work could be suitable for clinical laboratory practice. The diagnosis by smear showed sensitivity of 100% and specificity of 90%, but considering the classification of individual RBC images, the specificity raised up to 99.98%. This work was performed acquiring images using a camera attached to a microscope. The proposed approach could be also integrated in an automated image acquisition system, which would require an algorithm to find a suitable area of the smear for the proper assessment of RBC morphology. Furthermore, since some cases of malaria show very low levels of parasitaemia, the identification of infected RBC would be conditioned to the number of field images taken. Our system proved to be able to recognise malaria infection with parasitaemia values of 0.1% or higher.

Machine learning tools are becoming increasingly important in the field of laboratory medicine. This work developed a framework for the recognition of malaria in thin PBS digital images that improved specificity of previous approaches, being able to recognise other inclusions in RBC which are not related to malaria infection. The high sensitivity and specificity of our system reduces the number of smears to be reviewed by the clinical pathologist, improving the efficiency of the laboratory workflow.

Take home messages

- Thin blood smear microscopic examination is a low-expensive and easily accessible tool very valuable for malaria diagnosis. Image processing techniques may help to preclassify blood cells contributing to more standardised morphological analysis.
- The contribution of this paper is to provide an automatic recognition system able to discriminate parasitised RBC not only from normal RBC, but also from other erythrocyte inclusions, such as Howell-Jolly bodies, Pappenheimer bodies, basophilic stippling as well as platelets overlying RBC.
- The proposed system showed high sensitivity and specificity and it would be a practical tool to assist the pathologist in the diagnosis of malaria during the morphological peripheral blood examination.

Handling editor Mary Frances McMullin.

Contributors All authors have contributed to the development of the present work.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as online supplementary information.

ORCID iDs
Angel Molina http://orcid.org/0000-0002-9584-3646
Laura Boldú http://orcid.org/0000-0002-5162-3182
Anna Merino http://orcid.org/0000-0002-1889-8889

REFERENCES

1 World Health Organization. *World malaria report 2018. ISBN: 978-92-4-156565-3*.
2 Muñoz J, Rojo G, Ramírez G, *et al*. Diagnosis and treatment of imported malaria in Spain: recommendations from the malaria Working group of the Spanish Society of tropical medicine and international health (SEMTSI). *Enferm Infecc Microbiol Clin* 2015;33:e1–13.
3 World Health Organization. *Malaria microscopy quality assurance manual. version 2 (ISBN 978-92-4-154939-4)*, 2016.
4 Merino A. *Manual de Citología de Sangre Periférica*. Madrid: Acción Médica, 2005.
5 Bain BJ. Diagnosis from the blood smear. *N Engl J Med* 2005;353:498–507.
6 Constantino BT. Reporting and grading of abnormal red blood cell morphology. *Int J Lab Hematol* 2015;37:1–7.
7 Gutiérrez G, Merino A, Domingo A, *et al*. EQAS for peripheral blood morphology in Spain: a 6-year experience. *Int J Lab Hematol* 2008;30:460–6.
8 Díaz G, Manzanera A. Automatic analysis of microscopic images in hematological cytology applications. *In Clinical Technologies: Concepts, Methodologies, Tools and Applications* 2011:325–52.
9 Acevedo A, Alférez S, Merino A, *et al*. Recognition of peripheral blood cell images using convolutional neural networks. *Comput Methods Programs Biomed* 2019;180:105020.
10 Kratz A, Lee S-H, Zini G, *et al*. Digital morphology analyzers in hematology: ICSH review and recommendations. *Int J Lab Hematol* 2019;41:437–47.
11 Merino A, Puigví L, Boldú L, *et al*. Optimizing morphology through blood cell image analysis. *Int J Lab Hematol* 2018;40:54–61.
12 Rodellar J, Alférez S, Acevedo A, *et al*. Image processing and machine learning in the morphological analysis of blood cells. *Int J Lab Hematol* 2018;40:46–53.
13 Poostchi M, Silamut K, Maude RJ, *et al*. Image analysis and machine learning for detecting malaria. *Transl Res* 2018;194:36–55.
14 Ross NE, Pritchard CJ, Rubin DM, *et al*. Automated image processing method for the diagnosis and classification of malaria on thin blood smears. *Med Biol Eng Comput* 2006;44:427–36.
15 Moon S, Lee S, Kim H, *et al*. An image analysis algorithm for malaria parasite stage classification and viability quantification. *PLoS One* 2013;8:e61812.
16 Prasad K, Winter J, Bhat UM, *et al*. Image analysis approach for development of a decision support system for detection of malaria parasites in thin blood smear images. *J Digit Imaging* 2012;25:542–9.
17 Devi SS, Laskar RH, Sheikh SA. Hybrid classifier based life cycle stages analysis for malaria-infected erythrocyte using thin blood smear images. *Neural Comput Appl* 2018;29:217–35.
18 Gopakumar GP, Swetha M, Sai Siva G, *et al*. Convolutional neural network-based malaria diagnosis from focus stack of blood smear images acquired using custom-built slide scanner. *J Biophotonics* 2018;11:e201700003.
19 Das DK, Ghosh M, Pal M, *et al*. Machine learning approach for automated screening of malaria parasite using light microscopic images. *Micron* 2013;45:97–106.
20 Di Ruberto C, Dempster A, Khan S, *et al*. Analysis of infected blood cell images using morphological operators. *Image Vis Comput* 2002;20:133–46.
21 Diaz G, Gonzalez FA, Romero E. Infected cell identification in thin blood images based on color pixel classification: comparison and analysis. *Progress in Pattern Recognition, Image Analysis and Applications* 2007:812–21.
22 Das DK, Mukherjee R, Chakraborty C. Computational microscopic imaging for malaria parasite detection: a systematic review. *J Microsc* 2015;260:1–19.
23 Yang D, Subramanian G, Duan J, *et al*. A portable image-based cytometer for rapid malaria detection and quantification. *PLoS One* 2017;12:e0179161.
24 Linder N, Turkki R, Walliander M, *et al*. A malaria diagnostic tool based on computer vision screening and visualization of Plasmodium falciparum candidate areas in digitized blood smears. *PLoS One* 2014;9:e104855.

25 Díaz G, González FA, Romero E. A semi-automatic method for quantification and classification of erythrocytes infected with malaria parasites in microscopic images. *J Biomed Inform* 2009;42:296–307.
26 Tek FB, Dempster AG, Kale İzzet. Parasite detection and identification for automated thin blood film malaria diagnosis. *Comput Vis Image Underst* 2010;114:21–32.
27 Sheikhhosseini M, Rabbani H, Zekri M, *et al*. Automatic diagnosis of malaria based on complete circle-ellipse fitting search algorithm. *J Microsc* 2013;252:189–203.
28 Pan WD, Dong Y, Wu D. Classification of malaria-infected cells using deep convolutional neural networks, in machine learning: advanced techniques and emerging applications. *Intech Open* 2018;159.
29 Florin L, Maelegheer K, Muyldermans A, *et al*. Evaluation of the CellaVision DM96 advanced RBC application for screening and follow-up of malaria infection. *Diagn Microbiol Infect Dis* 2018;90:253–6.
30 Racsa LD, Gander RM, Southern PM, *et al*. Detection of intracellular parasites by use of the CellaVision DM96 analyzer during routine screening of peripheral blood smears. *J Clin Microbiol* 2015;53:167–71.
31 Piaton E, Fabre M, Goubin-Versini I, *et al*. [Technical recommendations and best practice guidelines for May-Grünwald-Giemsa staining: literature review and insights from the quality assurance]. *Ann Pathol* 2015;35:294–305.
32 Vives Corrons J-L, Albarède S, Flandrin G, *et al*. Guidelines for blood smear preparation and staining procedure for setting up an external quality assessment scheme for blood smear interpretation. Part I: control material. *Clin Chem Lab Med* 2004;42:922–6.
33 Alférez S, Merino A, Acevedo A, *et al*. Color clustering segmentation framework for image analysis of malignant lymphoid cells in peripheral blood. *Med Biol Eng Comput* 2019;57:1265–83.
34 van Griethuysen JJM, Fedorov A, Parmar C, *et al*. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 2017;77:e104–7.
35 Brown G, Pocock A, Zhao M-J, *et al*. Conditional likelihood maximisation: a unifying framework for information theoretic feature selection. *J Mach Learn Res* 2012;13:27–66.
36 Keng TB, De La Salle B, Bourner G, *et al*. Standardization of haematology critical results management in adults: an international Council for standardization in haematology, ICSH, survey and recommendations. *Int J Lab Hematol* 2016;38:457–71.
37 Le M-T, Bretschneider TR, Kuss C, *et al*. A novel semi-automatic image processing approach to determine Plasmodium falciparum parasitemia in Giemsa-stained thin blood smears. *BMC Cell Biol* 2008;9:15.
38 Purwar Y, Shah SL, Clarke G, *et al*. Automated and unsupervised detection of malarial parasites in microscopic images. *Malar J* 2011;10:364.
39 Somasekar J, Eswara Reddy B, Reddy BE. Segmentation of erythrocytes infected with malaria parasites for the diagnosis using microscopy imaging. *Comput Electr Eng* 2015;45:336–51.
40 Boldú L, Merino A, Alférez S, *et al*. Automatic recognition of different types of acute leukaemia in peripheral blood by image analysis. *J Clin Pathol* 2019;72:755–61.