



HAL
open science

Automatic patient-level recognition of four Plasmodium species on thin blood smear by a real-time detection transformer (RT-DETR) object detection algorithm: a proof-of-concept and evaluation

Emilie Guemas, Baptiste Routier, Théo Ghelfenstein-Ferreira, Camille Cordier, Sophie Hartuis, Bénédicte Marion, Sébastien Bertout, Emmanuelle Varlet-Marie, Damien Costa, Grégoire Pasquier

► To cite this version:

Emilie Guemas, Baptiste Routier, Théo Ghelfenstein-Ferreira, Camille Cordier, Sophie Hartuis, et al.. Automatic patient-level recognition of four Plasmodium species on thin blood smear by a real-time detection transformer (RT-DETR) object detection algorithm: a proof-of-concept and evaluation. Microbiology Spectrum, 2024, 12 (2), pp.e0144023. 10.1128/spectrum.01440-23 . hal-04607010

HAL Id: hal-04607010

<https://hal.science/hal-04607010>

Submitted on 10 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Automatic patient-level recognition of four *Plasmodium* species on thin blood smear by a real-time detection transformer (RT-DETR) object detection algorithm: a proof-of-concept and evaluation

Emilie Guemas,^{1,2} Baptiste Routier,³ Théo Ghelfenstein-Ferreira,⁴ Camille Cordier,⁵ Sophie Hartuis,⁶ Bénédicte Marion,^{7,8} Sébastien Bertout,⁹ Emmanuelle Varlet-Marie,^{7,8} Damien Costa,³ Grégoire Pasquier⁸

AUTHOR AFFILIATIONS See affiliation list on p. 10.

ABSTRACT Malaria remains a global health problem, with 247 million cases and 619,000 deaths in 2021. Diagnosis of *Plasmodium* species is important for administering the appropriate treatment. The gold-standard diagnosis for accurate species identification remains the thin blood smear. Nevertheless, this method is time-consuming and requires highly skilled and trained microscopists. To overcome these issues, new diagnostic tools based on deep learning are emerging. This study aimed to evaluate the performances of a real-time detection transformer (RT-DETR) object detection algorithm to discriminate *Plasmodium* species on thin blood smear images. The algorithm was trained and validated on a data set consisting in 24,720 images from 475 thin blood smears corresponding to 2,002,597 labels. Performances were calculated with a test data set of 4,508 images from 170 smears corresponding to 358,825 labels coming from six French university hospitals. At the patient level, the RT-DETR algorithm exhibited an overall accuracy of 79.4% (135/170) with a recall of 74% (40/54) and 81.9% (95/116) for negative and positive smears, respectively. Among *Plasmodium*-positive smears, the global accuracy was 82.7% (91/110) with a recall of 90% (38/42), 81.8% (18/22), and 76.1% (35/46) for *P. falciparum*, *P. malariae*, and *P. ovale/vivax*, respectively. The RT-DETR model achieved a World Health Organization (WHO) competence level 2 for species identification. Besides, the RT-DETR algorithm may be run in real-time on low-cost devices such as a smartphone and could be suitable for deployment in low-resource setting areas lacking microscopy experts.

IMPORTANCE Malaria remains a global health problem, with 247 million cases and 619,000 deaths in 2021. Diagnosis of *Plasmodium* species is important for administering the appropriate treatment. The gold-standard diagnosis for accurate species identification remains the thin blood smear. Nevertheless, this method is time-consuming and requires highly skilled and trained microscopists. To overcome these issues, new diagnostic tools based on deep learning are emerging. This study aimed to evaluate the performances of a real-time detection transformer (RT-DETR) object detection algorithm to discriminate *Plasmodium* species on thin blood smear images. Performances were calculated with a test data set of 4,508 images from 170 smears coming from six French university hospitals. The RT-DETR model achieved a World Health Organization (WHO) competence level 2 for species identification. Besides, the RT-DETR algorithm may be run in real-time on low-cost devices and could be suitable for deployment in low-resource setting areas.

Editor Kileen L. Shier, Quest Diagnostics Nichols Institute, Chantilly, Virginia, USA

Address correspondence to Grégoire Pasquier, g-pasquier@chu-montpellier.fr.

The authors declare no conflict of interest.

Received 4 April 2023

Accepted 15 November 2023

Published 3 January 2024

Copyright © 2024 Guemas et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

KEYWORDS *Plasmodium*, artificial intelligence, deep learning, machine learning, object detection, YOLO, RT-DETR, diagnosis, thin blood smear

Malaria is a vector-borne disease caused by a protozoan parasite belonging to the Apicomplexa taxon called *Plasmodium*. This disease is a global health problem, with 247 million cases and 619,000 deaths worldwide in 2021—increasing since 2019 (1). Over the 2 peak years of the pandemic (2020–2021), COVID-related disruptions led to about 13 million more malaria cases and 63,000 more malaria deaths (1). The diagnosis can be made by microscopy, rapid diagnostic tests, and molecular biology tests. Nevertheless, thin blood smears remain the gold standard technique, enabling species identification and parasitaemia calculation (2). Yet they require highly skilled and trained microscopists and do not prevent frequent misidentification (3). Species misdiagnosis may lead to inappropriate treatment, like chloroquine for *Plasmodium falciparum* or lack of anti-relapse treatment for *P. vivax* and *P. ovale* (4). The lack of microscopic diagnostic skills often gives unacceptable results (5), requiring training sessions (6), which are not always easy to implement.

Another recent approach could be to develop diagnostic tools based on deep learning, aiming at intra-erythrocytic *Plasmodium* detection on images of thick and thin blood smears. These are currently an expanding field of research in the microbiology field (7). The usual process of training neural network algorithms involves several steps: blood smear image acquisition; segmentation of smear images to obtain cell-sized images; and labeling these cell-sized images as infected or uninfected, for example, dividing the database of cell-sized images into an 80% training set and a 20% validation set, training the neural network, and evaluating the performance on a test set consisting of previously unseen images (8).

Many studies (9, 10) aimed to compare different deep learning algorithms, mainly convolutional neural networks (CNN) or their derivatives, using publicly available databases such as the National Institute of Health Malaria data set (11) or the Broad Bioimage Benchmark Collection (BBBC) (12). The more frequent drawbacks observed in these studies were the lack of a test data set, with the results calculated only on the validation data set; no patient-level results but only smear images or cell-sized results; too homogenous staining (not reflecting real diversity in routine practice); and issues related with the segmentation process. Segmentation of erythrocytes would be unsuccessful in real-life situations such as overlapping of red blood cells or differences in staining and slide preparation, especially if these cases are not included in the data set (13). To overcome these issues, object detection algorithm-based approaches including mask Regional-convolutional neural network (R-CNN) (14), faster R-CNN (12), and modified YOLOv3, v4, and v5 algorithms (15, 16) were developed, as they do not need a segmentation step and showed excellent results in terms of accuracy. Among them, YOLO-derived algorithms gave better results than the faster R-CNN (15, 17).

Current state of the art in malaria machine-learning-based detection is the EasyScan GO microscopy device, which reached the field evaluation stage using thick smears (18, 19). The same device was used to develop an algorithm framework including multiple CNNs for thin smears (20) and was evaluated on a 55-slide set from the World Health Organization (WHO) with 82.9% species identification accuracy on both positive and negative samples and 60.0% accuracy on positive samples exclusively (21). Nevertheless, EasyScan Go used $\times 400$ magnification images, whereas for precise species identification, microscopists used $\times 1,000$ magnification. Moreover, the EasyScan Go system runtime was 54.4 minutes per slide, which is above the threshold of 10 minutes per slide recommended by the WHO malaria microscopy quality assurance manual (21). Here, to try to overcome the aforementioned issues, a newly available real-time detection transformer (RT-DETR) object detection algorithm was trained and evaluated with a data set made of $\times 1,000$ magnification images taken from thin blood smears. Thus, this study aimed to evaluate and compare the accuracy and performances of the object detection machine-learning algorithm RT-DETR in the species detection of the four main

Plasmodium species. We also included in our study two less frequently encountered blood parasites: *Trypanosoma brucei* and *Babesia divergens*. The latter is intra-erythrocytic and can be confused with *Plasmodium* spp (22). Moreover, the performances of the RT-DETR algorithm were compared with those of the other two object detection algorithms: YOLOv5 and YOLOv8.

MATERIALS AND METHODS

Data collection

The training and validation data set included 24,720 pictures taken from 475 manually May Grunwald–Giemsa (MGG)-stained thin blood smears from the Montpellier University Hospital collection and for a smaller part from the Toulouse University Hospital collection. In Montpellier, the pictures were taken with a FlexCam C1 microscope camera (Leica) attached to a Leica DM 2000 microscope and Leica DF450C microscope camera adapted with a Leica DM2500 microscope at $\times 1,000$ magnification. Labeling of pictures was performed manually and then automatically with manual correction with a Computer Visual Annotation Tools (CVAT) free software. Nine categories of labels were used: white blood cells ($n = 3,338$), red blood cells ($n = 1,887,781$), platelets ($n = 48,520$), *Trypanosoma brucei* ($n = 2,773$), and red blood cells infected by *P. falciparum* ($n = 43,545$), *P. ovale* ($n = 4,651$), *P. vivax* ($n = 4,115$), *P. malariae* ($n = 2,732$), and *Babesia divergens* ($n = 5,142$).

The test data set included 4,508 pictures taken from 170 thin blood smears from the same number of patients from the Parasitology laboratories of University Hospitals of Montpellier, Toulouse, Rouen, Lille, Nantes, and Saint-Louis in Paris (Table 1). Among these 170 patients, 54 were not infected, including two patients with Howell–Jolly bodies, and 116 were infected with hematozoa. For each patient, between 20 and 30 photos were taken from one thin blood smear, with at least one hematozoan parasite per picture for infected patients.

Accurate species diagnosis was made by a senior parasitologist, and for recent smears, it was confirmed by specific PCR, either performed locally (Toulouse) or at the Malaria French National Reference Center (Montpellier, Saint Louis, Rouen, Lille, Nantes).

Algorithm training and validation

The training and validation data sets were split into 80% training and 20% validation to train the real-time object detection neural network algorithms RT-DETR (<https://arxiv.org/abs/2304.08069>), YOLOv5x (<https://github.com/ultralytics/yolov5>), and YOLOv8x (<https://github.com/ultralytics/ultralytics>) using the Pytorch framework. Images from the same patient contributed to either the training or validation data set, but not to both, to avoid bias. Parameters of algorithm training were as follows: image resolution 640 pp, epoch 25, patience 5, and batch size 32. The models were trained on an NVIDIA GeForce RTX3060 GPU with 12 GB of memory in 13.106 hours (14 epochs) for the RT-DETR model, in 14.976 hours (25 epochs) for the YOLOv5x model, and in 10.384 hours (23 epochs) for the YOLOv8x model. Overfitting was not observed during training of the models (Fig. S1 through S3).

Algorithm testing

The test data set images were resized at 16:9, and a zoom correction was applied to fit the images of the train/validation data set. Performances of the models were assessed at the label and patient level, thanks to the multicentric test data set. The label level corresponds to the raw data obtained directly from the object detection algorithm. Indeed, the object detection models work on a frame-by-frame basis. In each image, around a hundred objects are labeled, corresponding to uninfected red blood cells, platelets, and red blood cells infected by various parasites. The patient-level detection comes after this initial detection stage and consists of compiling the labels obtained

TABLE 1 Data set characteristics

Data set	Microscope	Camera	Smear number	Negative smears	<i>P. falciparum</i> smears	<i>P. ovale</i> smears	<i>P. malariae</i> smears	<i>P. vivax</i> smears	<i>Babesia divergens</i> smears	<i>Trypanosoma brucei</i> smears	
Training data set			403	57	176	52	48	40	8	22	
Montpellier	Leica Flexcam C1	Leica DM2500	379	57	171	47	43	35	6	20	
	Leica DM 2000	Leica Flexcam C1									
Toulouse	Nikon Eclipse E4000	Nikon DS-Fi2	24	0	5	5	5	5	2	2	
Validation data set	Leica Flexcam C1	Leica DM2500	72	19	15	9	5	17	2	5	
	Leica DM 2000	Leica Flexcam C1									
Montpellier			170	54	43	23	22	22	3	3	
Test data set			15	0	9	1	1	2	1	1	
Lille	Zeiss AXIO Imager.M2 4 slides Fluor/Color		65	54	11	0	0	0	0	0	
Montpellier	Leica Flexcam C1	Leica Flexcam C1	1	0	1	0	0	0	0	0	
Nantes	Leica DFC 280	Leica DM LB2	45	0	12	12	11	10	0	0	
Rouen	Olympus BX41	Olympus SC50	22	0	5	5	5	5	1	1	
Saint Louis	Olympus BX41	Olympus DP22	22	0	5	5	5	5	1	1	
Toulouse	Nikon Eclipse E4000	Nikon DS-Fi2	22	0	5	5	5	5	1	1	
Data set	Image number	Label number	White blood cell labels	Red blood cell labels	Platelet labels	<i>P. falciparum</i> labels	<i>P. ovale</i> labels	<i>P. malariae</i> labels	<i>P. vivax</i> labels	<i>Babesia divergens</i> labels	<i>Trypanosoma brucei</i> labels
Train Data set	20830	1677051	2949	1579822	42461	36241	3771	2164	3246	4262	2135
Montpellier	20209	1635166	2921	1540781	40961	35991	3635	2032	3093	3673	2079
Toulouse	621	41885	28	39041	1500	250	136	132	153	589	56
Validation dataset											638
Montpellier	3890	325546	389	307959	6059	7304	880	568	869	880	
Test dataset	4508	358825	416	342855	10384	2965	696	514	706	221	68
Lille	510	76027	24	73735	1222	665	52	33	187	77	32
Montpellier	2034	124226	243	117812	5440	684	1	0	46	0	0
Nantes	25	2882	0	2671	90	121	0	0	0	0	0
Rouen	981	107767	92	103148	2431	1269	263	279	285	0	0
Saint Louis	425	13570	33	12842	282	22	252	78	45	1	15
Toulouse	533	34353	24	32647	919	204	128	124	143	143	21

TABLE 2 Overall and per class results of the test dataset with the RT-DETR model^a

Class	Images	Labels	Precision	Recall	mAP@.5	mAP@.5:.95:	F1 score	MCC
All	4508	275240	0.686	0.669	0.638	0.596	0.677	0.633
WBC	4508	363	0.798	0.953	0.956	0.904	0.869	0.855
RBC	4508	261789	0.973	0.992	0.994	0.976	0.982	0.723
Platelets	4508	8581	0.895	0.956	0.979	0.9	0.924	0.852
<i>P. falciparum</i>	4508	2488	0.819	0.845	0.858	0.845	0.832	0.827
<i>P. ovale</i>	4508	530	0.417	0.353	0.199	0.193	0.382	0.38
<i>P. malariae</i>	4508	500	0.679	0.604	0.644	0.626	0.639	0.637
<i>P. vivax</i>	4508	672	0.376	0.284	0.15	0.143	0.324	0.326
<i>Babesia</i>	4508	250	0.385	0.32	0.203	0.197	0.35	0.34
<i>Trypanosoma brucei</i>	4508	67	0.827	0.716	0.764	0.579	0.768	0.756

^aMCC: Matthews correlation coefficient.

from several images belonging to the same smear to give a general result: uninfected, infected by this, or infected by that species.

At the label level, the performances for each label category were evaluated with precision = $\frac{TP}{TP + FP}$ with TP, true positive and FP, false positive; recall = $\frac{TP}{TP + FN}$

with FN, false negative; F1 score = $\frac{2(\text{precision} \times \text{recall})}{\text{precision} + \text{recall}}$; Matthews correlation coefficient = $\frac{(TN * TP - FP * FN)}{\sqrt{(TN + FN)(FP + TP)(TN + FP)(FN + TP)}}$ with TN, true negative; mean average precision at an intersection over union (IoU) of 0.5 (mAP@.5) and average mAP over different IoU thresholds, from 0.5 to 0.95, step 0.05 (mAP@.5:.95) (23) (Table 2).

At the patient level, the label confidence scores of each parasite were summed over the 20 to 30 pictures (Fig. 1), and the parasite with the highest sum of confidence scores was selected as the final diagnosis. Each label had a confidence score between 0 and 1. Results at patient level were presented in the form of a confusion matrix (Table 3) and were assessed in terms of accuracy (true predictions/ all predictions), precision, and recall.

RESULTS

Labels level

Out of the 4,508 images of the 170 smears in the test data set, 358,825 labels were generated by the model. At the label level, overall precision, recall, and mAP@.5 were 0.686, 0.669, and 0.638, respectively (Table 2). However, performance varied greatly by class. The mAP@.5 was greater than 0.95 for white blood cells, red blood cells, and platelets. The mAP@.5 was 0.858 for *P. falciparum*, 0.764 for *Trypanosoma brucei*, and 0.644 for *P. malariae*. This metric was lower for *P. ovale* (0.199) and *P. vivax* (0.15). The confusion matrix (Fig. 2) showed a significant mislabeling between the latter two parasites since 40% of the *P. ovale* were labeled as *P. vivax* and 35% of the *P. vivax* were labeled as *P. ovale*. The model poorly performed in detecting *Babesia divergens* parasites, with a precision of 0.385 and a recall of 0.32.

Parameters used for the confusion matrix were as follows: confidence score threshold equal to or greater than 0.25; IoU equal to or greater than 0.45; agnostic = True.

Patient level

At the patient level, the test data set included 170 thin blood smears from 170 patients, of which 54 were uninfected and 116 infected with either *P. falciparum*, *P. malariae*, *P. ovale*, *P. vivax*, *Babesia divergens*, or *Trypanosoma brucei* (Table 3). The overall accuracy with the six parasite class RT-DETR model was 68.2% (116/170). The recall was 100% (3/3) for *Trypanosoma brucei*, 90% (38/42) for *P. falciparum*, 74% (40/54) for negative smears, and 81.8% (18/22) for *P. malariae*. The recall was lower for *P. vivax* (26%, 6/23), *P. ovale* (43%, 10/23), and *Babesia divergens* (33%, 1/3). Misdiagnosis between *P. ovale* and *P. vivax* was

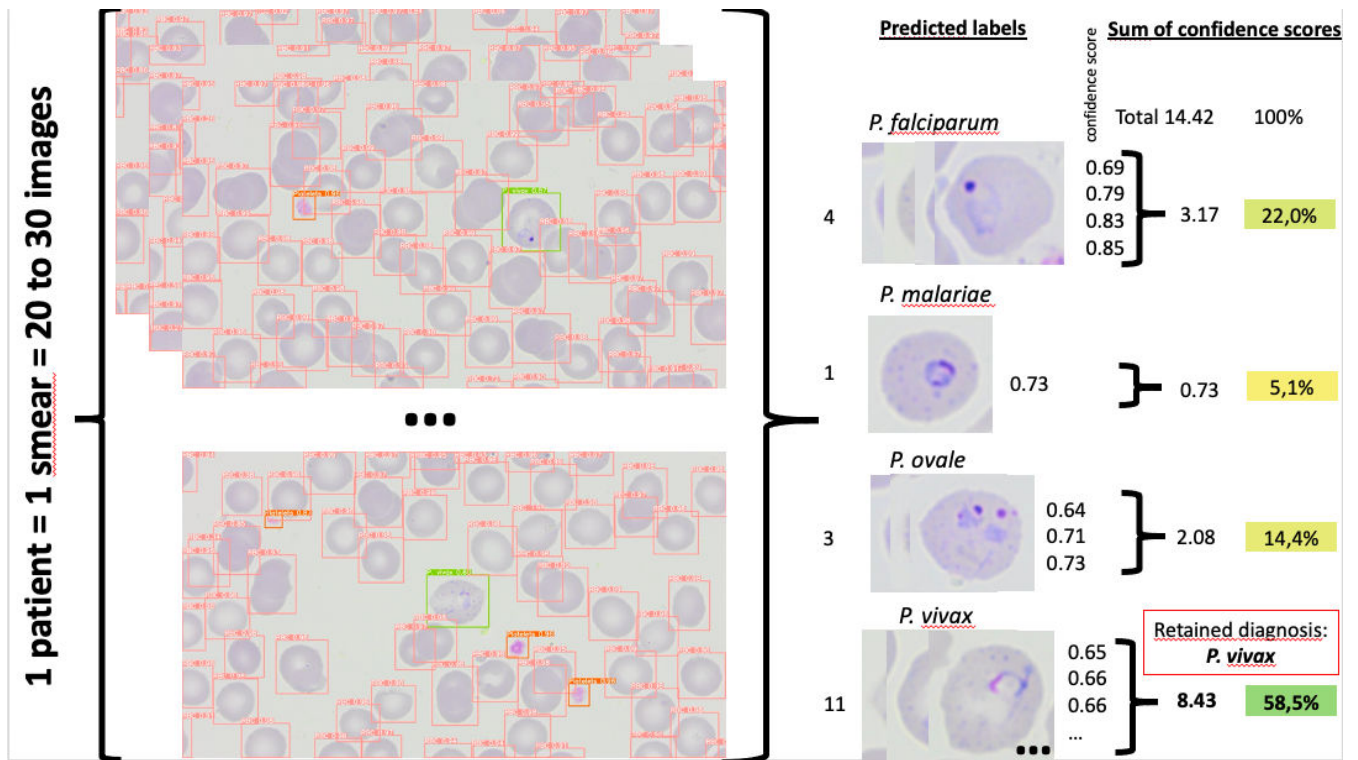


FIG 1 Determination of parasitic diagnosis at the patient level. Example of a smear from the test data set. *P. vivax* was selected as the final diagnosis because its sum of confidence scores is the highest (8.43 > 3.17 > 2.08 > 0.73).

TABLE 3 Results by class of the test data set at the patient level with the RT-DETR model^{a,b}

Predicted / True	<i>Babesia</i>	<i>P. falciparum</i>	<i>P. malariae</i>	<i>P. ovale</i>	<i>P. vivax</i>	Not infected	<i>Trypanosoma brucei</i>	Precision
<i>Babesia</i>	1	2			1			0.25
<i>P. falciparum</i>	1	38	4	2	8	3		0.68
<i>P. malariae</i>		1	18			10		0.62
<i>P. ovale</i>				10	8	1		0.53
<i>P. vivax</i>	1	1			11	6		0.32
Not infected						40		1
<i>Trypanosoma brucei</i>							3	1
Recall	0.33	0.90	0.82	0.43	0.26	0.74	1	
F1 score	0.29	0.78	0.71	0.48	0.29	0.85	1	
MCC	0.27	0.70	0.66	0.41	0.19	0.81	1	

^aMCC: Matthews correlation coefficient.

^bFigures in bold refer to metrics calculated from non-bold figures.

common, and among the 46 *P. ovale/vivax* smears, eight (17.4%) were *P. vivax* diagnosed as *P. ovale* and 11 (23.9%) were *P. ovale* diagnosed as *P. vivax*. Taking this into account, a five-parasite class model gathering *P. ovale* and *P. vivax* labels showed an overall accuracy of 79.4% (135/170) with a recall of 76.1% for the *P. ovale/vivax* class (35/46). For the classification of patients into infected and uninfected, the accuracy was 91.8% (156/170).

Algorithm comparison

The performance of the RT-DETR model was compared with that of the YOLOv5x model and of the YOLOv8x model (Table 4). At the label level, YOLOv8x seems to be slightly better than the other models, with an mAP@.5 of 0.727 (Fig. S14) vs 0.67 for the YOLOv5x (Fig. S25) and 0.638 for RT-DETR (Table 2). At the patient level, the overall accuracy of the

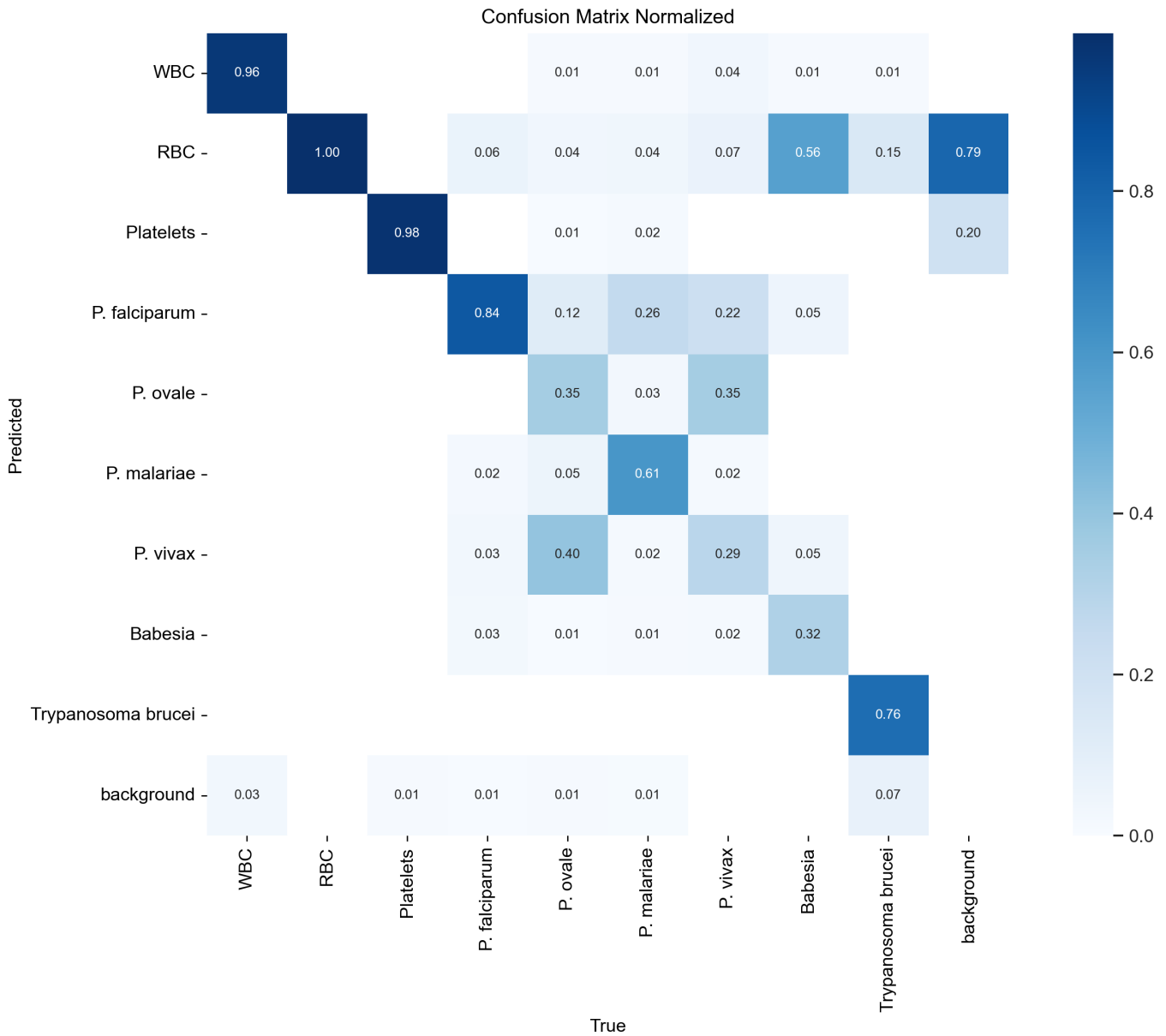


FIG 2 Confusion matrix with test data set labels of the RT-DETR model.

three models was similar: 62.8% (116/170) for the RT-DETR and the YOLOv5x model (Table S3) and 67% (114/170) for the YOLOv8x model (Table S4).

DISCUSSION

To our knowledge, the RT-DETR model is the first algorithm aiming to identify five hematozoan parasites on thin blood smear images. With a large test data set of 170 thin blood smears from six different centers, the six-parasite class model exhibited an overall accuracy of 68.2% (116/170), while the five-parasite class model, which combines *P. ovale* and *P. vivax*, showed an overall accuracy of 79.4% (135/170) at the patient level.

In respect to the results by parasite class at the patient level, the recall for *Trypanosoma brucei* was 100% (3/3), which was expected due to its easily identifiable extra-erythrocytic nature. The recall for *P. falciparum* was 90% (38/42). Four *P. falciparum* smears were misdiagnosed as *P. malariae* (one smear), *P. vivax* (one smear), and *B. divergens* (two smears). Old trophozoites with Maurer’s clefts were particularly prone to being misdiagnosed as a non-*falciparum* species.

TABLE 4 Comparison of patient-level results of three algorithm models^a

		YOLOv5x	YOLOv8x	RT-DETR
Accuracy, number of	Six-parasite class model	116 (68.2)	114 (67)	116 (68.2)
Correct diagnosis /170 (%)	Five-parasite class model	137 (80.6)	129 (75.3)	135 (79.4)
	Four-parasite class model	139 (81.8)	137 (80.6)	135 (79.4)
	Two-parasite class model	160 (94.1)	160 (94.1)	156 (91.8)
Accuracy, number of	Four- <i>Plasmodium</i> class model	66 (60)	66 (60)	72 (65.5)
Correct diagnosis /110 (%)	Three- <i>Plasmodium</i> class model	87 (79.1)	81 (73.6)	91 (82.7)
	Two- <i>Plasmodium</i> class model	89 (81)	89 (81)	91 (82.7)
<i>Babesia divergens</i>	Precision	0	0.33	0.25
	Recall	0	0.33	0.33
	F1-score	0	0.33	0.29
<i>P. falciparum</i>	Precision	0.63	0.73	0.68
	Recall	0.86	0.69	0.90
	F1-score	0.73	0.71	0.78
<i>P. malariae</i>	Precision	0.93	0.61	0.62
	Recall	0.59	0.5	0.82
	F1-score	0.72	0.55	0.71
<i>P. ovale</i>	Precision	0.40	0.41	0.53
	Recall	0.74	0.91	0.43
	F1-score	0.52	0.57	0.48
<i>P. vivax</i>	Precision	0	0.56	0.32
	Recall	0	0.21	0.26
	F1-score	0	0.31	0.29
Not infected	Precision	0.98	1	1
	Recall	1	0.81	0.74
	F1-score	0.9	0.90	0.85
<i>Trypanosoma brucei</i>	Precision	0.6	0.69	1
	Recall	1	1	1
	F1-score	0.75	0.75	1

^aSix-parasite class model (*T. brucei*, *B. divergens*, *P. falciparum*, *P. malariae*, *P. ovale*, and *P. vivax*, none infected). Five-parasite class model (*T. brucei*, *B. divergens*, *P. falciparum*, *P. malariae*, and *P. ovale/vivax*, none infected). Four-parasite class model (*T. brucei*, *B. divergens*, *P. falciparum*, and *P. malariae/ovale/vivax*, none infected). Two-parasite class model (*T. brucei*, intraerythrocytic parasites, none infected). Four-*Plasmodium* class model (*P. falciparum*, *P. malariae*, *P. ovale*, and *P. vivax*). Three-*Plasmodium* class model (*P. falciparum*, *P. malariae*, and *P. ovale/vivax*). Two-*Plasmodium* class model (*P. falciparum* and *P. malariae/ovale/vivax*).

The recall for *P. ovale* and *P. vivax* was 43% (10/23) and 26% (6/23), respectively, showing a high rate of confusion between these two species. Indeed, 19 out of 46 *P. vivax/ovale* smears were misdiagnosed as *P. vivax* for *P. ovale* or vice versa. This is coherent with the difficulties encountered by microscopists in routine practice to distinguish both species without geographic priors or rapid diagnostic test information (3). Nevertheless, this error has no therapeutic or clinical impact on patients since the treatment is similar for both species. Thus, grouping *P. ovale* and *P. vivax* in a single class achieved a recall of 76.1% (35/46).

The recall for *P. malariae* was 81.8% (18/22). Four *P. malariae* smears were misclassified as *P. falciparum* due to a high rate of young ring-form trophozoites. A new class could have been trained for *P. malariae* late parasitic stages (equatorial-like bands; rosette schizonts) which are species-specific patterns, but the number of corresponding labels was too small in the data set.

The recall for *Babesia divergens* was only 33% (1/3)%. This may be due to the lack of *Babesia* smears included in the train/validation data set. The recall of negative smears was 74% (40/54). Forty-one out of 100,708 labels (0.04%) corresponding to 14 out of 54 negative smears (25.9%) of the test data set were misclassified as positive. Ten of them were mislabeled as *P. malariae*, three as *P. falciparum*, and one as *P. ovale* (Fig. S5).

A previous study (21) using the EasyScan GO device showed 60% accuracy in identifying the four major *Plasmodium* species on a set of 15 positive smears. This is to

be compared with the presented RT-DETR model, which exhibited an overall accuracy of 65.5% (72/110) on 110 *Plasmodium*-positive smears and an 82.7% (91/110) accuracy with the model pooling *P. ovale* and *P. vivax*. The latter model achieved a WHO competence level 2 for species identification (24). For species identification, the presented algorithm could be of great help for inexperienced microscopists and shows advantages over both loop-mediated isothermal amplification (LAMP), which does not discriminate between *Plasmodium* species, and rapid diagnostic tests, which have a very low sensitivity for *P. malariae* and *P. ovale*.

YOLO was described as more performant than other object detection algorithms (R-CNN, SSD) using publicly available malaria data set (15, 17), and RT-DETR was found more performant than YOLO models on the COCO (Common Objects in Context) data set (<https://arxiv.org/abs/2304.08069>). The RT-DETR algorithm gets rid of the computationally expensive segmentation step used with the CNN to obtain multiple cropped images with a unique red blood cell from a microscope field of view image. Besides, the RT-DETR algorithm can be run in real-time on low-cost devices such as a smartphone with a microscope adapter (25). Such a system could be adapted to improve parasite detection for inexperienced microscopists and to alert them to forms suspected of being *Plasmodium* parasites. A later step will therefore require adapting this RT-DETR algorithm to a device that can be used in the field. Mobile applications have been developed using either CNN, R-CNN such as Malaria Screener (26), or PlasmoCount (27). Beyond manual microscopy, other systems could be considered to reach partial or full automatization such as a robotic handed microscope (28) or hardware slide scan (29).

This study acknowledged some limitations. It was not designed to assess parasite detection and quantification performance of the RT-DETR model but only for species identification purposes. Parasitemia assessment could not be done on our test data set since the images were not captured randomly but focused on parasite forms. Although the test data set included images from six different parasitology laboratories, previously unseen staining artifacts are likely to induce algorithm errors. As for species identification, the parasitic stage should also be considered. Indeed, although late-stage trophozoites, schizonts, and gametocytes are more specific for species identification, they are less abundant than ring and young trophozoites. For this purpose of parasite stage recognition, an unsupervised machine learning approach might be interesting (30). Besides, performances might be improved by coupling the RT-DETR algorithm with expert rules including geographic priors, parasitemia, and rapid diagnostic tests results.

To our knowledge, this study is the first attempt to predict malaria species from thin blood smear images using an RT-DETR object detection algorithm. This algorithm allows accurate real-time detection of parasites on easily affordable devices as smartphones mounted on a microscope. Further studies should aim to develop and evaluate this cell phone application, particularly in malaria endemic countries where trained microscopists are not in sufficient numbers.

ACKNOWLEDGMENTS

Our special thanks go to Denis Vigouroux, IT engineer, for his help with the YOLO algorithm. We would also like to thank the members of the Hematology (Théo Lannay, Julien Houssier, Jennifer Russello, and Patricia Aguilar Martinez) and Bacteriology (Lucas Bonzon, Charlie Zins, and Sylvain Godreuil) laboratories of the Montpellier University Hospital and the Parasitology-Mycology laboratory of the Pharmacy of Pharmacy at University of Montpellier (Donika Krasteva and Virginie Bellet), for allowing us to use microscopic pictures from their respective laboratories. We also wish to thank Bernard Carcy for providing us with *Babesia* smears. We are also grateful to Patrick Bastien for help in revising the manuscript.

This work was supported by the Montpellier University Hospital.

G.P. designed the study, took smear pictures, annotated database images, conducted statistical analysis, and drafted and revised the manuscript. E.G. took smear pictures and drafted and revised the manuscript. B.R. designed the study, took smear pictures, and

revised the manuscript. T.G.F., C.C., S.H., and B.M. took smear pictures. S.B. revised the manuscript. E.V. designed the study and revised the manuscript. D.C. designed the study and revised the manuscript. All authors have read and approved the final manuscript.

No conflict of interest to be reported for G.P., E.G., B.R., T.G.F., S.H., C.C., B.M., E.V., S.B., and D.C.

AUTHOR AFFILIATIONS

¹Department of Parasitology and Mycology, Academic Hospital (CHU) of Toulouse, Toulouse, France

²Toulouse Institute for Infectious and Inflammatory Diseases (Infinity), CNRS UMR5051, INSERM UMR1291, UPS, Toulouse, France

³Laboratory of Parasitology-Mycology, EA7510 ESCAPE, University Hospital of Rouen, University of Rouen Normandie, Normandie, France

⁴Université de Paris Cité, Laboratoire de Parasitologie-Mycologie, Groupe Hospitalier Saint-Louis-Lariboisière-Fernand-Widal, Assistance Publique-Hôpitaux de Paris (AP-HP), Paris, France

⁵Laboratory of Parasitology-Mycology, INSERM U1285, Unité de Glycobiologie Structurale et Fonctionnelle (CNRS UMR 8576), University Hospital (CHU) of Lille, University of Lille, Lille, France

⁶Nantes University, Academic Hospital (CHU) of Nantes, Cibles et Médicaments des Infections et de l'Immunité, IICiMed, UR1155, Nantes, France

⁷Department of Physical Chemistry and Biophysics, Academic Hospital (CHU) of Montpellier, University of Montpellier, National Reference Centre (CNR) for Paludism, Montpellier, France

⁸Department of Parasitology/Mycology, Academic Hospital (CHU) of Montpellier, University of Montpellier, National Reference Centre (CNR) for Paludism, Montpellier, France

⁹Laboratory of Parasitology/Mycology, UMI 233 TransVIHMI, University of Montpellier, IRD, INSERM U1175, Montpellier, France

AUTHOR ORCID*s*

Emilie Guemas  <http://orcid.org/0000-0002-0168-7575>

Grégoire Pasquier  <http://orcid.org/0000-0002-9152-5149>

AUTHOR CONTRIBUTIONS

Emilie Guemas, Conceptualization, Data curation, Validation, Writing – original draft, Writing – review and editing | Baptiste Routier, Data curation, Formal analysis, Writing – review and editing | Théo Ghelfenstein-Ferreira, Data curation, Resources, Writing – review and editing | Camille Cordier, Data curation, Investigation, Writing – review and editing | Sophie Hartuis, Data curation, Investigation, Writing – review and editing | Bénédicte Marion, Data curation, Investigation | Sébastien Bertout, Investigation, Supervision, Writing – review and editing | Emmanuelle Varlet-Marie, Conceptualization, Data curation, Investigation, Writing – review and editing | Damien Costa, Conceptualization, Data curation, Investigation, Supervision, Writing – review and editing | Grégoire Pasquier, Conceptualization, Data curation, Formal analysis, Methodology, Software, Supervision, Validation, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

Data and models are available on Zenodo ([DOI: 10.5281/zenodo.8358829](https://doi.org/10.5281/zenodo.8358829)).

ADDITIONAL FILES

The following material is available [online](#).

Supplemental Material

Supplemental Figures (Spectrum01440-S0001.docx). Figures S1 to S6.
Supplemental Tables (Spectrum01440-S0002.docx). Tables S1 to S5.

REFERENCES

- World Malaria Report. 2022. Geneva : World Health Organization, CC BY-NC-SA 3.0 IGO. Available from: <https://www.who.int/teams/global-malaria-programme>
- White NJ, Pukrittayakamee S, Hien TT, Faiz MA, Mokuolu OA, Dondorp AM. 2014. Malaria. The Lancet 383:723–735. [https://doi.org/10.1016/S0140-6736\(13\)60024-0](https://doi.org/10.1016/S0140-6736(13)60024-0)
- Kotepui M, Masangkay FR, Kotepui KU, De Jesus Milanez G. 2020. Misidentification of *Plasmodium ovale* as *plasmodium vivax* malaria by a microscopic method: a meta-analysis of confirmed *P. ovale* cases. Sci Rep 10:21807. <https://doi.org/10.1038/s41598-020-78691-7>
- Barber BE, William T, Grigg MJ, Yeo TW, Anstey NM. 2013. Limitations of microscopy to differentiate plasmodium species in a region co-endemic for *Plasmodium falciparum*, *Plasmodium vivax* and *Plasmodium knowlesi*. Malar J 12:8. <https://doi.org/10.1186/1475-2875-12-8>
- Obare P, Ogutu B, Adams M, Odera JS, Lilley K, Dosoo D, Adhiambo C, Owusu-Agyei S, Binka F, Wanja E, Johnson J. 2013. Misclassification of plasmodium infections by conventional microscopy and the impact of remedial training on the proficiency of laboratory technicians in species identification. Malar J 12:113. <https://doi.org/10.1186/1475-2875-12-113>
- Edson DC, Glick T, Massey LD. 2010. Detection and identification of malaria parasites: a review of proficiency test results and laboratory practices. Lab Med 41:719–723. <https://doi.org/10.1309/LMOKC4BEYHGDSZCU>
- Peiffer-Smadja N, Dellièrre S, Rodriguez C, Birgand G, Lescure F-X, Fourati S, Ruppé E. 2020. Machine learning in the clinical microbiology laboratory: has the time come for routine practice? Clin Microbiol Infect 26:1300–1309. <https://doi.org/10.1016/j.cmi.2020.02.006>
- Poostchi M, Silamut K, Maude RJ, Jaeger S, Thoma G. 2018. Image analysis and machine learning for detecting malaria. Transl Res 194:36–55. <https://doi.org/10.1016/j.trsl.2017.12.004>
- Jameela T, Athotha K, Singh N, Gunjan VK, Kahali S. 2022. Deep learning and transfer learning for malaria detection. Comput Intell Neurosci 2022:2221728. <https://doi.org/10.1155/2022/2221728>
- Khan SH, Shah NS, Nuzhat R, Majid A, Alquhayz H, Khan A. 2022. Malaria parasite classification framework using a novel channel squeezed and boosted CNN. Microscopy 71:271–282. <https://doi.org/10.1093/jmicro/dfac027>
- Rajaraman S, Antani SK, Poostchi M, Silamut K, Hossain MA, Maude RJ, Jaeger S, Thoma GR. 2018. Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. PeerJ 6:e4568. <https://doi.org/10.7717/peerj.4568>
- Hung J, Carpenter AE. 2017. Applying faster R-CNN for object detection on malaria images 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); Honolulu, HI, USA: p 808–813 IEEE. <https://doi.org/10.1109/cvprw.2017.112>
- Depto DS, Rahman S, Hosen MM, Akter MS, Reme TR, Rahman A, Zunair H, Rahman MS, Mahdy MRC. 2021. Automatic segmentation of blood cells from microscopic slides: a comparative analysis. Tissue Cell 73:101653. <https://doi.org/10.1016/j.tice.2021.101653>
- Loh DR, Yong WX, Yapeter J, Subburaj K, Chandramohanadas R. 2021. A deep learning approach to the screening of malaria infection: automated and rapid cell counting, object detection and instance segmentation using mask R-CNN. Comput Med Imaging Graph 88:101845. <https://doi.org/10.1016/j.compmedimag.2020.101845>
- Abdurahman F, Fante KA, Aliy M. 2021. Malaria parasite detection in thick blood smear microscopic images using modified YOLOV3 and YOLOV4 models. BMC Bioinform 22:112. <https://doi.org/10.1186/s12859-021-04036-4>
- Krishnadas P, Chadaga K, Sampathila N, Rao S, S. SK, Prabhu S. 2022. Classification of malaria using object detection models. Informatics 9:76. <https://doi.org/10.3390/informatics9040076>
- Rocha M, Claro M, Neto L, Aires K, Machado V, Veras R. 2023. Malaria parasites detection and identification using object detectors based on deep neural networks: a wide comparative analysis. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization 11:351–368. <https://doi.org/10.1080/21681163.2022.2111715>
- Das D, Vongpromek R, Assawariyathipat T, Srinamon K, Kennon K, Stepniewska K, Ghose A, et al. 2022. Field evaluation of the diagnostic performance of easyscan GO: a digital malaria microscopy device based on machine-learning. Malar J 21:122. <https://doi.org/10.1186/s12936-022-04146-1>
- Torres K, Bachman CM, Delahunt CB, Alarcon Baldeon J, Alava F, Gamboa Vilela D, Proux S, Mehanian C, McGuire SK, Thompson CM, Ostbye T, Hu L, Jaiswal MS, Hunt VM, Bell D. 2018. Automated microscopy for routine malaria diagnosis: a field comparison on giemsa-stained blood films in Peru. Malar J 17:339. <https://doi.org/10.1186/s12936-018-2493-0>
- Delahunt CB, Jaiswal MS, Horning MP, Janko S, Thompson CM, Kulhare S, et al. 2022. Fully-automated patient-level malaria assessment on field-prepared thin blood film microscopy images, including supplementary information. arXiv. <https://doi.org/10.48550/arXiv.1908.01901>
- Horning MP, Delahunt CB, Bachman CM, Luchavez J, Luna C, Hu L, Jaiswal MS, Thompson CM, Kulhare S, Janko S, Wilson BK, Ostbye T, et al. 2021. Performance of a fully-automated system on a WHO malaria microscopy evaluation slide set. Malar J 20:110. <https://doi.org/10.1186/s12936-021-03631-3>
- Vannier E, Krause PJ. 2012. Human babesiosis. N Engl J Med 366:2397–2407. <https://doi.org/10.1056/NEJMra1202018>
- Padilla R, Netto SL, da Silva EAB. 2020. A survey on performance Metrics for object-detection Algorithms 2020 International Conference on Systems, Signals and Image Processing (IWSSIP); Niterói, Brazil: , p 237–242 IEEE, Brazil. <https://doi.org/10.1109/IWSSIP48289.2020.9145130>
2016. World health organization. malaria microscopy quality assurance manual. version 2. Geneva:World Health Organization. Available from: <https://apps.who.int/iris/handle/10665/204266>
- Chibuta S, Acar AC. 2020. Real-time malaria parasite screening in thick blood smears for low-resource setting. J Digit Imaging 33:763–775. <https://doi.org/10.1007/s10278-019-00284-2>
- Yu H, Yang F, Rajaraman S, Ersoy I, Moallem G, Poostchi M, Palaniappan K, Antani S, Maude RJ, Jaeger S. 2020. Malaria screener: a smartphone application for automated malaria screening. BMC Infect Dis 20:825. <https://doi.org/10.1186/s12879-020-05453-1>
- Davidson MS, Andradi-Brown C, Yahiya S, Chmielewski J, O'Donnell AJ, Gurung P, Jenning MD, Prommana P, Andrew DW, Petter M, Uthaipibull C, Boyle MJ, et al. 2021. Automated detection and staging of malaria parasites from cytological smears using convolutional neural networks. Biol Imaging 1:e2. <https://doi.org/10.1017/S2633903X21000015>
- Vijayalakshmi A, Rajesh Kanna B. 2020. Deep learning approach to detect malaria from microscopic images. Multimed Tools Appl 79:15297–15317. <https://doi.org/10.1007/s11042-019-7162-y>
- Yoon J, Kwon JA, Yoon SY, Jang WS, Yang DJ, Nam J, Lim CS. 2019. Diagnostic performance of cellavision DM96 for *Plasmodium vivax* and *Plasmodium falciparum* screening in peripheral blood smears. Acta Trop 193:7–11. <https://doi.org/10.1016/j.actatropica.2019.02.009>
- Li S, Du Z, Meng X, Zhang Y. 2021. Multi-stage malaria parasite recognition by deep learning. Gigascience 10:giab040. <https://doi.org/10.1093/gigascience/giab040>